EFFECTS OF PROCESSING VARIATIONS AND ENDURANCE STRESS ON THE MNOS NONVOLATILE MEMORY DEVICE

BY

ALOYSIUS ANSCRUTHA MARINO DE ALMEIDA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1986

To

my wife,

my children,

and the memory of my parents

"...Oars alone can ne'er prevail

To reach the distant coast,

The breath of heav'n must swell the sail,

Or all the toil is lost."

Cowper

ACKNOWLEDGEMENTS

I wish to thank the chairman of my Ph.D. committee, Dr. Sheng Li, for the guidance and inspiration he has given me throughout my Ph.D. work. I also wish to thank the other committee members, Drs. Dorothea Burk, Gijs Bosman, Paul Holloway and Arun Varma for being in my committee. I am grateful to Ben Stevens of the Office of Medical Education for providing me with a graduate assistantship and for allowing me to use his office facilities to prepare this manuscript. I appreciate the technical assistance given by Dr. Ken Ports and William O'Mara of Harris Semiconductor in the fabrication of devices and in exercising them.

I wish to express my deep gratitude to my parents, the late Millany Aloysius and the late Sylnet Aloysius, who, beyond the gift of life, have endowed me with a will to persevere against all odds. Last, but by no means the least, I wish to thank my wife Linda and my sons Shehan and Rukshan, who have borne with equanimity the vicissitudes of my Ph.D. sojourn. Linda's support and forbearance and her help in preparing this manuscript are invaluable.

CONTENTS

ACKNOWLEDGEMENTS iii

ABBREVIATIONS vii

SYMBOLS ix

ABSTRACT xvi

CHAPTER Page

I INTRODUCTION 1

II DEVICE PHYSICS AND CONDUCTION PROCESSES 7

- 2.1 Energy Band Diagrams and C-V Curves 7
- 2.2 Physical Parameters 14
- 2.3 Admittance Parameters 22
- 2.4 Conduction Processes 24

III MNOS ADMITTANCE AND MECHANICS OF OPERATION 32

- 3.1 Derivation of the Small Signal Equivalent Circuit 32
- 3.2 Equivalent Circuit with Reference to the Energy Band Diagram 35
- 3.3 Carrier Response 38
- 3.4 C-V and G-V Characteristics 39
- 3.5 G-ω Characteristics 43
- 3.6 Energy Level Behavior 44

IV DEVICE FABRICATION, EXPERIMENTS AND ANALYSIS 45

- 4.1 Device Fabrication and Features 45

4.2 Measurements 46

4.3 QSCV Analysis and Results 48

4.4 Model Verification 56

V TRAPS, CHARGES AND THEIR ORIGINS 61

5.1 Chemical Composition and Structure 61

5.1.1 Measuring Techniques 61

5.1.2 Silicon Nitride 62

5.1.3 Silicon Dioxide and Si/SiO$_2$ and Oxide/Nitride Interfaces 63

5.2 Nitride Traps, Oxide/Nitride Interface Traps and Their Origins 66

5.3 Si/SiO$_2$ Interface States, Silicon Bulk Traps and Their Origins 68

5.4 Charges 73

5.5 Measured Si/SiO$_2$ Interface-State Density Curves and Their Dependence on Processing Variations 77

VI RETENTION 79

6.1 Introduction 79

6.2 Effects of Processing Variations on the Retention Characteristics 81

6.3 Measured Retention Characteristics and Their Dependence on Processing Variations 83

VII ENDURANCE 93

7.1 Introduction 93

7.2 Degradation Models 96

7.3 Experimental Results of Degradation Studies 99

7.4 Quantitative Illustration of Degradation Effects 116

VIII CONCLUSIONS AND RECOMMENDATIONS 120

8.1 Conclusions 120

8.2 Recommendations 122

REFERENCES 123

APPENDIX

A PHYSICAL CONSTANTS 132

B MNOS QUASI-STATIC ANALYSIS PROGRAM 133

C MNOS C-V CURVE SIMULATION PROGRAM 137

BIOGRAPHICAL SKETCH 139

ABBREVIATIONS

at.%	Atomic Percentage
B-T	Bias - Temperature
CCDLTS	Constant Capacitance DLTS
C-V	Capacitance - Voltage
DCS	Dichlorosilane (SiH_2Cl_2)
DLTS	Deep Level Transient Spectroscopy
ESR	Electron Spin Resonance
G-R	Generation - Recombination
G-V	Conductance - Voltage
G-ω	Conductance - Angular Frequency
HREM	High Resolution Electron Microscopy
IRA	Infra Red Absorption
LPCVD	Low Pressure Chemical Vapor Deposition
MIR	Multiple Internal Reflection
MNOS	Metal Nitride Oxide Semiconductor
NRA	Nuclear Reaction Analysis
QSCV	Quasi-Static C-V
RHEED	Reflection High-Energy Electron Diffraction
RBS	Rutherford Backscattering Spectrometry
SAES	Sputter Auger Electron Spectroscopy
Si 2p PS	Silicon 2p Photoelectron Spectra

TE	Thermal Emission
TSC	Thermally Stimulated Current
VAP	Valance Alternate Pair
XPS	X-ray Photoelectron Spectroscopy
1st V/dec	V'_{TH} decay rate per decade of time before the break point
2nd V/dec	V'_{TH} decay rate per decade of time after the break point

SYMBOLS

A	Device area (cm^2)
A_1	Area of the n-region (cm^2)
C_D	Depletion capacitance of the n-region (pF)
C_I	Insulator capacitance (pF)
C_{In}	Insulator capacitance between the n-region and the gate (pF)
C_{INV}	Inversion capacitance of the n-region (pF)
C_{Ip}	Insulator capacitance between the p^+-grid and the gate (pF)
C_M	High frequency measured capacitance (pF)
C_{min}	Minimum high frequency measured capacitance (pF)
C_{M0}	Quasi-static measured capacitance (pF)
C_N	Nitride capacitance (pF)
C_{0x}	Oxide capacitance (pF)
C_p	Si/SiO_2 interface-state capacitance for majority carriers (pF)
C_p'	Si/SiO_2 interface-state capacitance for minority carriers (pF)
C_{pn}	p^+-n junction capacitance (pF)
C_{p0}	Quasi-static value of C_p (pF)
C_{sc}	Silicon space charge capacitance (pF)
C_{sc0}	Quasi-static value of C_{sc} (pF)
E_c	Conduction band edge energy level in the silicon bulk (eV)
E_{cN}	Nitride conduction band edge energy level when no field is applied (eV)

E'_{cN}	Nitride conduction band edge energy level when a field is applied (eV)
E_{cs}	Conduction band edge energy level at the silicon surface (eV)
E_D	Donor energy level in the silicon (eV)
E_F	Fermi energy level in the silicon (eV)
E_{Fm}	Fermi energy level in the metal (eV)
E_g	Bandgap energy of silicon (eV)
E_i	Midgap energy level in the silicon bulk (eV)
E_{is}	Midgap energy level at the silicon surface (eV)
E_{ss}	Energy level of an Si/SiO_2 interface state (eV)
E_t	Nitride trap energy level measured from the nitride conduction band edge (eV)
E_T	Silicon bulk trap energy level (eV)
E_v	Valence band edge energy level in the silicon bulk (eV)
E_{vs}	Valence band edge energy level at the silicon surface (eV)
F_N	Nitride field (V/cm)
F_{Ox}	Oxide field (V/cm)
F_s	Silicon surface field (V/cm)
$F_{Si}(x)$	Silicon field at a distance x from the Si/SiO_2 interface (V/cm)
G	Silicon space charge loss conductance (μS)
G_d	Silicon minority carrier diffusion and drift conductance (μS)
G_{gr}	G-R conductance of silicon bulk traps (μS)
G_{grs}	G-R conductance of Si/SiO_2 interface states (μS)
G_p	Si/SiO_2 interface-state conductance for majority carriers (μS)
G'_p	Si/SiO_2 interface-state conductance for minority carriers (μS)
G_{p0}	Quasi-static value of G_p (μS)
G_s	Series conductance (μS)

i_G	Total small signal current leaving the gate (A)
i_{G1}	Small signal current leaving the gate due to n-region charges only (A)
i_{pn}	Small signal current flowing into the silicon due to the p^+-n junction charge (A)
i_{sc}	Small signal current flowing into the silicon due to the silicon space charge (A)
i_{ss}	Small signal current flowing into the silicon due to the Si/SiO_2 interface-state charge (A)
J_A	Oxide Fowler-Nordheim tunneling current density (A/cm^2)
J_b	Oxide back-tunneling current density (A/cm^2)
J_B	Oxide direct tunneling-to-traps current density (A/cm^2)
J_C	Oxide direct band-to-band tunneling current density (A/cm^2)
J_N	Total nitride current density (A/cm^2)
J_{Ox}	Total oxide current density (A/cm^2)
J_1	Nitride Poole-Frenkel current density (A/cm^2)
J_2	Nitride field (or tunnel) emission current density (A/cm^2)
J_3	Nitride Ohmic current density (A/cm^2)
k	Boltzmann constant ($J/^oK$)
N_A	Acceptor density of the p^+-grid (cm^{-3})
N_D	Donor density of the n-region (cm^{-3})
n(x)	Nitride trapped charge density at a distance x from the oxide/nitride interface (cm^{-3})
n_i	Intrinsic carrier density of silicon (cm^{-3})
n_o	Equilibrium electron density in the silicon bulk (cm^{-3})
N_o	Nitride trapped charge density at the oxide/nitride interface (cm^{-3})
n_s	Electron density at the silicon surface (cm^{-3})
N_{ss}	Si/SiO_2 interface-state density ($eV^{-1}cm^{-2}$)

N_t	Nitride trap density (cm^{-3})
p_o	Equilibrium hole density in the silicon bulk (cm^{-3})
p_s	Hole density at the silicon surface (cm^{-3})
q	Electronic charge (C)
Q	Localized charge in the oxide (C)
Q_d	Silicon depletion layer charge (C)
Q_{di}	Value of Q_d at onset of inversion (C)
Q_{dmax}	Value of Q_d at onset of strong inversion (C)
Q_I	Initial stored charge (C)
Q_M	Total charge on the metal gate (C)
Q_{M1}	Charge on the gate due to silicon charges only (C)
Q_n	Silicon inversion layer charge (C)
Q_{OxN}	Stored charge at the oxide/nitride interface (C)
Q_{pn}	p^+-n junction charge (C)
Q_{sc}	Silicon space charge (C)
Q_{Si}	Total silicon charge (C)
Q_{ss}	Si/SiO_2 interface-state charge (C)
Q_T	Total stored charge (C)
r	Constant of V_{TH} decay
t	Time (s)
T	Temperature (^oK)
t_d	Time of onset of initial decay (s)
T_N	Thickness of the nitride layer (Å)
T_{Ox}	Thickness of the oxide layer (Å)
V_{bi}	Built-in potential across the p^+-n junction (V)
V_{FB}	Flatband voltage (V)

V_G	d.c. component of V_G^* (i.e. bias voltage) (V)
V_G^*	Total gate voltage (V)
V_G'	a.c. component of V_G^* (V)
V_{Gm}	Amplitude of V_G' (V)
V_{Gs}	d.c. component of V_{Gs}^* (V)
V_{Gs}^*	Total voltage across the series conductance (V)
V_{Gs}'	a.c. component of V_{Gs}^* (V)
V_{Gsm}	Amplitude of V_{Gs}' (V)
V_n	Volume of the nitride (cm^3)
V_N	Voltage across the nitride (V)
V_{Ox}	Voltage across the oxide (V)
v_{th}	Thermal velocity of electrons (cm/s)
V_{TH}	Threshold voltage (V)
V_{TH}'	A measure of V_{TH} used for retention measurements (V)
V_{TI}	Initial value of V_{TH} (V)
x	Distance from the Si/SiO_2 interface or the oxide/nitride interface (Å)
x_c	Distance of the nitride charge centroid from the oxide/nitride interface (Å)
x_d	Silicon depletion width (Å)
x_{di}	Value of x_d at onset of inversion (Å)
x_{dmax}	Value of x_d at onset of strong inversion (Å)
x_t	Effective back-tunneling distance from the charge centroid of Q_I (Å)
Y	Admittance of the series combination of C_{In} and Y_1 (μS)
Y_{ss}	Si/SiO_2 interface-state admittance (μS)
Y_T	Total device admittance (μS)
Y_1	Admittance of the n-region (includes G where appropriate) (μS)

β	Reciprocal of the thermal voltage (V^{-1})
β_{PF}	Poole-Frenkel constant ($V^{1/2}cm^{1/2}$)
ΔE_{ae}	Activation energy of thermally excited electrons which give rise to Ohmic current (eV)
ΔN_t	Change in the nitride trap density due to degradation (cm^{-3})
ΔQ_{ss}	Increase in the Si/SiO_2 interface-state charge due to degradation (C)
ΔQ_T	Change in the nitride stored charge due to degradation (C)
ΔV_{sc}	Contribution from the change in the nitride stored charge to $\Delta V'_{TH}$ (V)
ΔV_{ss}	Contribution from the increase in the Si/SiO_2 interface-state density to $\Delta V'_{TH}$ (V)
ΔV_{TH}	Shift in V_{TH} due to degradation (V)
$\Delta V'_{TH}$	Shift in V'_{TH} due to degradation (V)
$\Delta \phi_{Ba}$	Potential barrier lowering due to image force effect (for Fowler-Nordheim current) (V)
$\Delta \phi_{PF}$	Poole-Frenkel barrier lowering (for Poole-Frenkel current) (V)
ε_N	Permittivity of the nitride (F/cm)
ε_{Ox}	Permittivity of the oxide (F/cm)
ε_{Si}	Permittivity of the silicon (F/cm)
λ	Mean free path for trapping of charges in the nitride (Å)
$\rho_N(x)$	Volume density of the nitride charge at a distance x from the oxide/nitride interface (C/cm^3)
$\rho_{Ox}(x)$	Volume density of the oxide charge at a distance x from the Si/SiO_2 interface (C/cm^3)
σ_n	Majority carrier capture cross-section of a continuum of Si/SiO_2 interface states (cm^2)
τ_n	Majority carrier time constant of a continuum of Si/SiO_2 interface states (s)
τ_0	Free carrier lifetime (s)
ϕ_B	Potential difference between the Fermi and midgap energy levels in the silicon bulk (V)

ϕ_{Ba}	Potential barrier height for Fowler-Nordheim injection (V)
ϕ_{M}	Metal work function (eV)
ϕ_{MS}	Metal-semiconductor work function difference (eV)
$\phi_{N}(x)$	Nitride potential at a distance x from the oxide/nitride interface (V)
$\phi_{Ox}(x)$	Oxide potential at a distance x from the Si/SiO$_2$ interface (V)
ϕ_{OxN}	Oxide-nitride work function difference (eV)
ϕ_{OxS}	Oxide-silicon work function difference (eV)
ϕ_{s}	d.c. component of ϕ_s^* (V)
ϕ_s^*	Total silicon surface potential (V)
ϕ_s'	a.c. component of ϕ_s^* (V)
ϕ_{sm}	Amplitude of ϕ_s' (V)
$\phi_{Si}(x)$	Silicon potential at a distance x from the Si/SiO$_2$ interface (V)
χ	Electron affinity of silicon (eV)
ω	angular frequency (rads/s)
ω_m	Frequency at which the maximum of G_p/ω occurs (rads/s)
+Q	Denotes positive stored charge
-Q	Denotes negative stored charge

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EFFECTS OF PROCESSING VARIATIONS AND ENDURANCE STRESS ON THE MNOS NONVOLATILE MEMORY DEVICE

By

Aloysius Anscrutha Marino de Almeida

May 1986

Chairman: Sheng-San Li
Major Department: Electrical Engineering

The p^+-gridded MNOS capacitor, which enables the simulation of the operational characteristics of the MNOS FET, was used to study the effects of processing parameter variations and endurance stress on the Si/SiO_2 interface-state density and retention characteristics of the MNOS. A model was derived theoretically for this device, and based on this model, a FORTRAN program was written to extract the device parameters from the measured C-V data using the QSCV technique. A second FORTRAN program was developed to simulate the experimental C-V curves using the measured device parameters. The simulation of the high frequency C-V curves in the accumulation region necessitated a slight modification of the basic model. Four processing variations were used in this study.

The measured Si/SiO_2 interface-state density distribution across the silicon bandgap has a general U shape, modified by humps in the upper or the lower half of the bandgap. The interface-state density minimum values for the four experimental devices ranged from $2x10^{11}$ to 10^{12} $eV^{-1}cm^{-2}$. The differences in these characteristics were interpreted in terms of the different chemical compositions and structures resulting from the variations in the processing conditions.

The retention characteristics were measured for the four devices, and these characteristics and their differences were interpreted in terms of the different chemical compositions and structures. The effects of increased temperature on the retention characteristics were studied, and the differences in these characteristics were used to infer relative nitride trap depths in the four devices.

The effects of Write/Erase cycling and bias-temperature stress on the Si/SiO_2 interface-state density and the retention characteristics of the four devices were studied. The effects of Write/Erase cycling on the surface potential versus gate voltage curve were also studied. The higher and broader Si/SiO_2 interface-state density curves resulting from endurance stress were interpreted in terms of different degradation models. These models along with the chemical compositions and structures arising from the different processing conditions were used to explain the effects of endurance stress on the retention characteristics.

CHAPTER I
INTRODUCTION

The Metal-Nitride-Oxide-Semiconductor (MNOS) device is used as a non-volatile memory element. The threshold voltage, which defines the memory state, can be altered from its original value to either of two stable state values, one more positive (0) and one more negative (1), by the application of a voltage pulse of the appropriate polarity and sufficient magnitude. The principle of the MNOS device lies in the charge accumulation near the oxide/nitride interface during charging due to the difference in the conductivities of the two insulators and trapping of this charge at the oxide/nitride interface and in the nitride. The trapped charge undergoes very slow back-tunneling through the oxide and slow redistribution into the nitride during retention and the application of low Read voltages. The schematics of an MNOS capacitor and a drain-source protected (DSP) MNOS FET are shown in Figure 1.1 [1]. In the latter, the thick oxide layer adjacent to the source and the drain protect the thin oxide memory region from high electric fields and crystal defects in the source and the drain.

Because its memory state is electrically alterable, the MNOS is used in EAROM's. Being non-volatile, it conserves power and is useful in power down situations. Hence, it has military applications such as in radar signal processors and guidance and control systems. It has an endurance of up to about 10^{10} Write/Erase cycles, which enables its

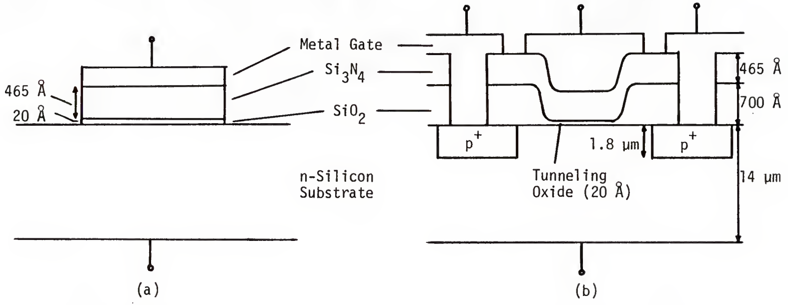


Fig.1.1 The schematics of (a) an MNOS capacitor and (b) a drain-source protected MNOS FET.

use in RAM's such as the block-oriented RAM (BORAM) [1]. MNOS is not suitable in very rapidly reprogrammed memories such as computer main memories, where the limit of endurance may be exceeded in a very short time. However, it can be used in computer secondary memories. It has applications also in airborne and spacecraft recorders and stored program systems. It has a cumulative retention time of up to ten years and a retention capability up to about 150°C [1,2].

The MNOS was developed as a result of the efforts to resolve the ion drift problem in MOS devices due primarily to contaminant sodium ions. This ion drift causes the threshold voltage to shift and is therefore undesirable in memory elements. Silicon nitride, which has a higher density than silicon dioxide, causes the mobility of sodium ions in it to be lower. The use of a thick nitride layer along with a thin oxide layer avoids the ion drift problem and inversion failure of thick oxide MOS and seals the surface against oxygen, water vapor and other contaminants. Silicon dioxide, on the other hand, is permeable to moisture and radiation. The double dielectric layer also inhibits time dependent dielectric breakdown (TDDB) associated with MOS devices, since the probability that the defects in the two layers would be lined up to facilitate this failure is remote. The SiO_2 layer serves as a charge trapping barrier and reduces (passivates) the large number of states at the silicon surface due to the abrupt termination of the periodic lattice.

The Metal-Nitride-Semiconductor (MNS) device has been considered as a possible memory device, but it has been shown to have a very significant threshold voltage hysterisis. This is due to charge trapping at the nitride/silicon interface, which has a high state density. Also, the absence of the oxide layer gives poorer retention in this device than in the MNOS.

Interface-doped MNOS devices have been studied [2]. Since most of the charge is trapped at the doped oxide/nitride interface, these devices require smaller and shorter Write pulses. This enables the use of a thicker oxide to obtain improved retention. However, because of reduced endurance, these devices have limited applicability.

A recent development in the Metal-Insulator-Semiconductor (MIS) family is the Metal-Oxide-Nitride-Oxide-Semiconductor (MONOS) [3]. The top oxide layer enables the use of a nitride layer as thin as 50 Å, so that Write/Erase voltages of only about ± 6 V are required. Thus, MONOS can be used in MOS LSI circuits, where the use of MNOS, which needs voltages of ± 25 V, is precluded. The top oxide blocks hole injection from the gate, thus preventing the creation of traps in the oxide and at the Si/SiO_2 interface. Therefore, the MONOS is expected to have better retention and endurance characteristics than the MNOS. However, this has yet to be verified, as the MONOS is still in an early stage of development.

The Si/SiO_2 interface is considered to play a very important role in retention and device degradation. Therefore, it has been studied extensively using the Quasi-Static Capacitance (QSCV), Conductance, Thermally Stimulated Current (TSC), Deep Level Transient Spectroscopy (DLTS) and Constant Capacitance DLTS (CCDLTS) techniques. Of these, the QSCV technique is the simplest and most economical and is therefore very useful when a large number of measurements are involved such as in a study of the effects of varying the processing parameters and of Write/Erase cycling. Also, as relative values are of primary interest in such studies, errors tend to cancel out. The QSCV technique can be made very accurate by modelling the experimental device accurately and recognizing the bias range within which only deterministic parameters play a significant role.

Charging and discharging mechanisms and basic device physics of the MNOS device are now well established. Although various mechanisms for charge trapping, charge decay and degradation have been proposed, these processes are not yet fully understood. The most recent research efforts have been directed at studying the chemical structure and composition of the important regions of the device and the effects of varying the processing parameters on device performance, retention and degradation.

A p^+-grid diffusion on the n-silicon substrate enables the use of an MNOS capacitor to simulate the operational characteristics of an MNOS FET, while avoiding the multiple masking operations required for the latter. It facilitates the rapid injection and removal of minority carriers into and from the silicon nitride and the silicon inversion region. Thus, Write/Erase speeds comparable to those of MNOS FET's are achieved. Being readily available, the minority carriers respond to a.c. signals of considerably higher frequencies than in MOS or regular MNOS devices.

The gridded structure has been used for charge transfer and retention studies [4], but no attempt has been made to model it. In this research, a model is derived for the p^+-gridded MNOS device, and the QSCV technique [5] is adapted for this structure to extract the interface-state density distribution in the silicon energy bandgap from C-V measurements. The dependence of retention and endurance on the interface-state density and of retention on temperature and the effects of processing parameter variations on the interface-state density, retention and endurance are studied. The results are interpreted in terms of chemical compositions and structures.

In Chapter II, device physics and conduction processes are described. The admittance parameters of the MNOS and the mechanisms underlying the behavior of these parameters during voltage and frequency sweeps are discussed in Chapter III. The fabrication process, the salient features of the device and the measuring techniques used in this research are described in Chapter IV. Also in this chapter, a slightly modified version of the theoretical device model is used to simulate the experimental C-V curves. The traps and the charges in the different regions of the device, their chemical origins and the influences of different processing conditions on these traps and charges are discussed in Chapter V. In Chapter VI, the retention characteristics and their temperature and processing variation dependences are described. Degradation effects and the influences of processing variations on these effects are studied in Chapter VII. Conclusions of this research and recommendations for future work are given in Chapter VIII.

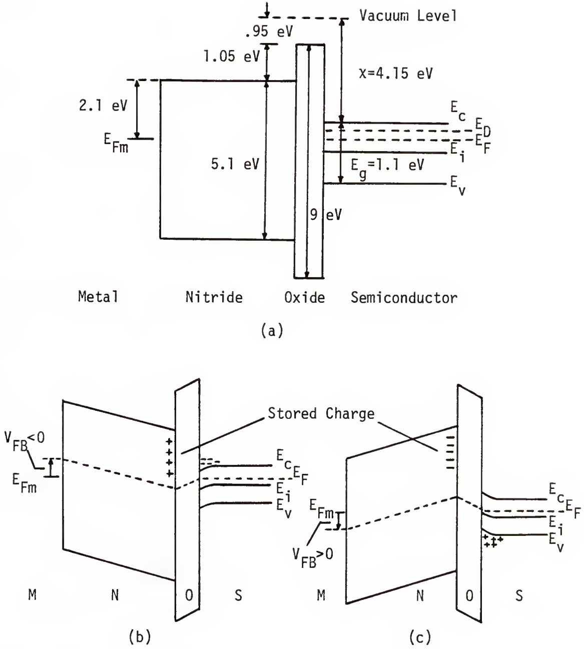
In Chapter II, charges, capacitances and conductances in expressions have per unit area values. The QSCV analysis of MNOS is similar to that of MOS, provided that the C-V sweep does not induce memory charges. Values of constants are listed in Appendix A.

CHAPTER II
DEVICE PHYSICS AND CONDUCTION PROCESSES

2.1 Energy Band Diagrams and C-V Curves

Figure 2.1 shows the energy band diagram for n-silicon MNOS at zero bias for the ideal case and for positive and negative trapped charges, where E_C, E_V, E_i, E_F, E_D, E_g and E_{Fm} are respectively the conduction band edge, valence band edge and midgap energy levels in the silicon bulk, the Fermi and donor energy levels of silicon, the bandgap energy of silicon and the Fermi energy level of the metal; χ is the electron affinity of silicon; and V_{FB} is the flatband voltage. Figure 2.2 shows the high frequency C-V curves for the three cases of Figure 2.1, where C_M is the measured capacitance, and V_G and V_{TH} are the gate and threshold voltages respectively. The factors that cause deviations from the ideal energy band diagram are the stored memory charge; the metal-semiconductor work function difference; thermal oxides, which give rise to deviation in the oxide energy bands due to inhibition of carrier conduction [6]; and thin oxides, which give rise to non-abrupt band edges due to oxynitride formation [7].

Figure 2.3 shows the low and high frequency C-V curves of the n-silicon MNOS capacitor, depicting the three characteristic charge regions. Figure 2.4 shows the energy band diagrams corresponding to these three regions. The following definitions in terms of the silicon



(a)

(b)

(c)

Fig.2.1 Energy band diagram of n-silicon MNOS at zero bias for (a) the ideal case, (b) positive stored charge and (c) negative stored charge.

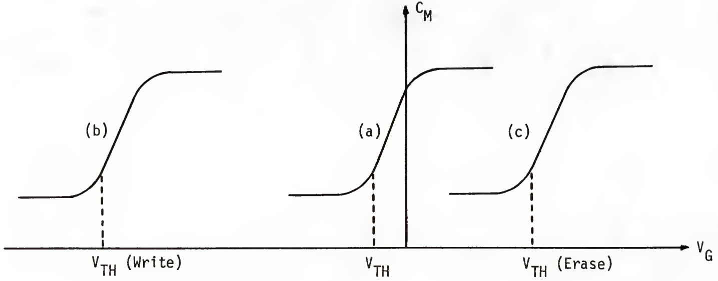


Fig.2.2 High frequency C-V curves of n-silicon MNOS for (a) the ideal case, (b) positive stored charge and (c) negative stored charge.

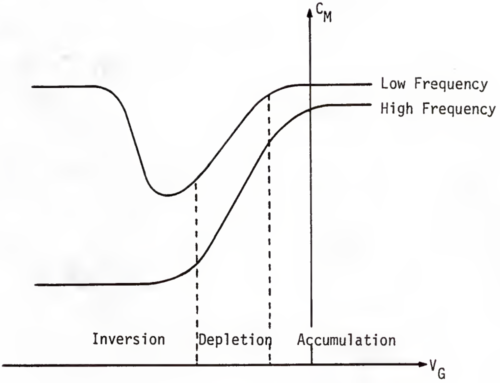
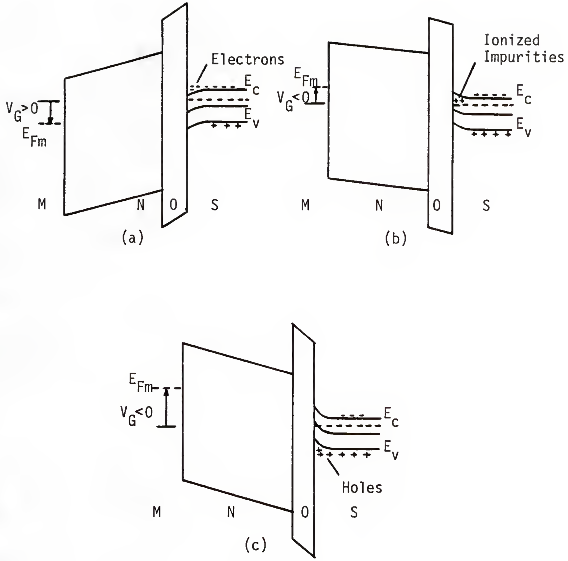


Fig.2.3 Low and high frequency C-V curves of n-silicon MNOS, depicting the accumulation, depletion and inversion regions.



(a)

(b)

(c)

Fig.2.4 Energy band diagram of n-silicon MNOS in (a) accumulation, (b) depletion and (c) inversion.

surface potential ϕ_s and the silicon surface carrier densities are used to physically distinguish the characteristic charge regions:

(a) Accumulation: $\phi_s > 0$; the surface majority carrier density $n_s >$ the donor density N_D.

(b) Flat-bands: $\phi_s = 0$; $n_s = N_D$.

(c) Depletion: $\phi_s < 0$; the surface is depleted of majority carriers; the silicon charge is due to ionized donors only; $n_s < N_D$.

(d) Onset of inversion: $\phi_s = -\phi_B$; where ϕ_B is the bulk potential (i.e. the difference between the Fermi and midgap energy levels in the silicon bulk; $\phi_B > 0$ for n-silicon device); n_s = the surface minority carrier density p_s; i.e. the surface is intrinsic.

(e) Weak inversion: $\phi_s < -\phi_B$; $p_s > n_s$.

(f) Onset of strong inversion: $\phi_s = -2\phi_B$; the depletion width saturates; $p_s = N_D$.

(g) Strong inversion: $\phi_s < -2\phi_B$; $p_s > N_D$.

Figure 2.5 shows the actual energy band diagram, where ϕ_M, ϕ_{OxN}, ϕ_{OxS}, V_N and V_{Ox} are respectively the metal work function, the oxide-nitride and Si-SiO$_2$ work function differences and the voltages across the nitride and the oxide; and E_{cs}, E_{is} and E_{vs} are respectively the conduction band edge, midgap and valence band edge energy levels at the silicon surface. The convention used is that the positive potentials point downwards, and vice-versa. At zero gate bias, the Fermi levels in the metal and the silicon equalize by electron transfer from the silicon to the metal through some external conductive path. The electrons are drawn from the silicon surface, causing depletion and

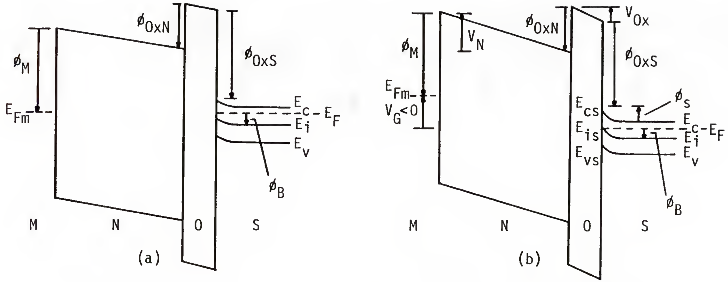


Fig.2.5 Actual energy band diagram of n-silicon MNOS, showing potentials at (a) zero bias and (b) an inversion bias.

band-bending. The negative charge of the electrons at the gate is balanced by the positive ionic charge at the silicon surface.

2.2 Physical Parameters

Figure 2.6 shows the charge distribution in the MNOS device, where Q_M, Q_{OxN}, Q_{ss}, Q_n and Q_d are the gate, the oxide/nitride interface, the Si/SiO$_2$ interface-state, the inversion layer and the depletion layer charges respectively; $\rho_N(x)$ is the nitride charge density at a distance x from the oxide/nitride interface; and T_{Ox} and T_N are the oxide and nitride thicknesses respectively.

From the depletion approximation and the integral of Poisson's equation, the electric field in the silicon depletion layer at a distance x from the Si/SiO$_2$ interface is obtained as [8]

$$F_{Si}(x) = (qN_D/\varepsilon_{Si})(x-x_d) \tag{2.1}$$

where q is the electron charge, ε_{Si} is the silicon permittivity, and x_d is the depletion layer width. Applying Gauss's Law, the oxide and nitride electric fields are obtained from the silicon surface field F_s as respectively

$$F_{Ox} = (\varepsilon_{Si}/\varepsilon_{Ox})F_s \tag{2.2}$$

$$F_N = (\varepsilon_{Ox}/\varepsilon_N)F_{Ox} \tag{2.3}$$

where ε_{Ox} and ε_N are the oxide and nitride permittivities respectively. The silicon, oxide and nitride potentials at a distance x from the Si/SiO$_2$ interface are obtained by integrating the corresponding

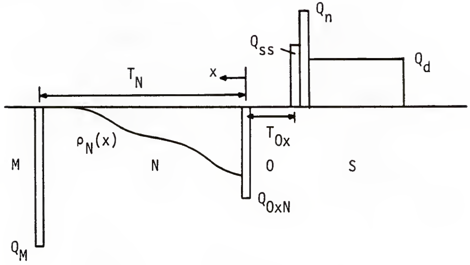


Fig.2.6 Complete charge distribution in the n-silicon MNOS capacitor in inversion (for negative stored charge).

electric fields as respectively

$$\phi_{Si}(x) = (qN_D/2\varepsilon_{Si})(x_d-x)^2 \tag{2.4}$$

$$\phi_{Ox}(x) = F_{Ox}x + \phi_s \tag{2.5}$$

$$\phi_N(x) = F_N\cdot(x-T_{Ox}) + \phi_{Ox}\Big|_{x=-T_{Ox}} \tag{2.6}$$

The ideal electric field and potential distributions in the MNOS device are shown in Figure 2.7.

From the charge neutrality condition, Kirchhoff's voltage law, Faraday's law and Gauss's law, the gate voltage is obtained as [2]

$$V_G = - Q_{Si}/C_I - (1-x_c/T_N)Q_T/C_N + \phi_{MS} + \phi_s \tag{2.7}$$

where $Q_{Si} = Q_{sc} + Q_{ss}$ is the total silicon charge; Q_{sc} is the silicon space charge, which is equal to Q_d in depletion and Q_n+Q_d in inversion; C_I is the total insulator capacitance; x_c is the distance of the nitride charge centroid from the oxide/nitride interface, given by

$$x_c = (1/Q_T)\int_0^{T_N} x\rho_N(x)dx \tag{2.8}$$

Q_T is the total stored charge, given by

$$Q_T = Q_{OxN} + \int_0^{T_N} \rho_N(x)dx \tag{2.9}$$

C_N is the nitride capacitance; and ϕ_{MS} is the metal-semiconductor work function difference, given by

$$\phi_{MS} = \phi_M + \phi_{OxN} - (\phi_{OxS}-\phi_B+E_g/2) \tag{2.10}$$

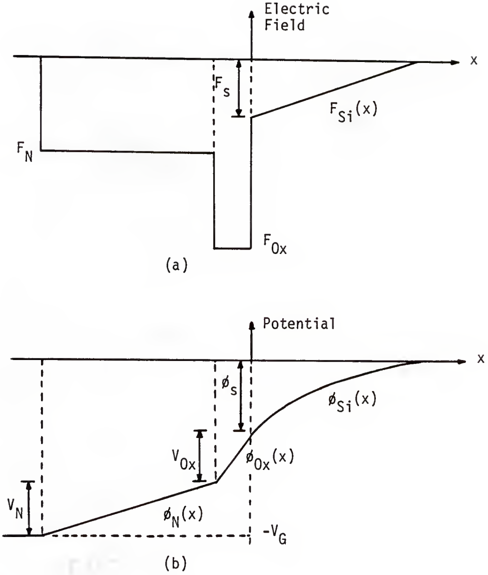


Fig.2.7 Ideal (a) electric field and (b) potential distributions of the n-silicon MNOS capacitor in depletion.

If a localized charge Q was present in the oxide at a distance x from the Si/SiO$_2$ interface, the charge quantity $(1-xC_I/\varepsilon_{Ox})Q$ must be added to Q_{Si} in (2.7). The effect of this charge is minimum when $x = T_{Ox}$ (i.e. at the oxide/nitride interface) and is maximum when $x = 0$ (i.e. at the Si/SiO$_2$ interface). If a distributed oxide charge was present with a density $\rho_{Ox}(x)$ at a distance x from the Si/SiO$_2$ interface, the following must be added to the right side of (2.7):

$$-(1/C_I)\int_0^{T_{Ox}} \rho_{Ox}(x)dx + (1/\varepsilon_{Ox})\int_0^{T_{Ox}} x\rho_{Ox}(x)dx$$

From (2.1), the electric field at the n-silicon surface is obtained as

$$F_s = -qN_Dx_d/\varepsilon_{Si} \tag{2.11}$$

where x_d is obtained by integrating (2.1) from the depletion layer edge to the Si/SiO$_2$ interface and is given as

$$x_d = [-2\varepsilon_{Si}\phi_s/(qN_D)]^{1/2} \qquad \phi_s < 0 \tag{2.12}$$

In practice, a correction factor of kT/q, where k is the Boltzmann constant, and T is the absolute temperature, must be added to the numerical value of ϕ_s to account for the non-abrupt termination in the depletion layer charge distribution. From Gauss's law, the depletion layer charge is

$$Q_d = -\varepsilon_{Si}F_s \tag{2.13}$$

Substituting from (2.11),

$$Q_d = qN_Dx_d \tag{2.14}$$

Substituting from (2.12),

$$Q_d = (-2q\varepsilon_{Si}N_D\phi_s)^{1/2} \qquad \phi_s < 0 \qquad (2.15)$$

From (2.12), the values of the depletion layer width at onset of inversion and onset of strong inversion are obtained as respectively

$$x_{di} = [2\varepsilon_{Si}\phi_B/(qN_D)]^{1/2} \qquad \phi_B > 0 \qquad (2.16)$$

$$x_{dmax} = [4\varepsilon_{Si}\phi_B/(qN_D)]^{1/2} \qquad (2.17)$$

These values along with (2.14) give the respective charge quantities Q_{di} and Q_{dmax}. The depletion width saturates to x_{dmax} at onset of strong inversion. This is due to the silicon being screened by the inversion layer, which prevents any further ionization of the donor atoms. Thereafter, even a slight increase in the electric field results in a large increase in the inversion charge.

The flatband voltage V_{FB} is obtained from (2.7) along with the conditions that ϕ_s and Q_{sc} are zero. Then, the electron and hole densities are uniform throughout the silicon.

$$V_{FB} = -\,Q_{ss}/C_I - (1-x_c/T_N)Q_T/C_N + \phi_{MS} \qquad (2.18)$$

If doping is not uniform, Q_{sc} may not be zero at flatbands. A physical interpretation of V_{FB} is shown in Figure 2.1. The threshold voltage, that is the voltage at onset of inversion, is also obtained from (2.7) along with the condition that Q_n is zero. Then, the electron and hole densities are equal at every point in the silicon.

$$V_{TH} = -\,(Q_{ss}+Q_{di})/C_I - (1-x_c/T_N)Q_T/C_N + \phi_{MS} - \phi_B \qquad (2.19)$$

Assuming a constant Q_{ss} (true for a given capacitance value), the only cause for a change in V_{FB} or V_{TH}, which gives a translation of the C-V curve along the voltage axis, is a change in the stored charge, since Q_{di}, ϕ_{MS} and ϕ_B are all constant. The change in both V_{FB} and V_{TH} is equal to the change in $-(1-x_c/T_N)Q_T/C_N$. Non-uniform doping and donor type interface traps contribute negative components to V_{FB} and V_{TH}, while acceptor type interface traps contribute a positive component. Sze [9] reports that the change in the threshold voltage increases linearly with logarithm of the Write time for a given pulse height, with a tendency to saturate at large times, and that the logarithm of the Write time required to obtain a given shift in the threshold voltage decreases linearly with pulse height.

Solving the Poisson's equation, the silicon surface electric field is obtained as [9]

$$F_s = \pm\,[2qn_0/(\varepsilon_{Si}\beta)]^{1/2}[(p_0/n_0)(e^{-\beta\phi_s}+\beta\phi_s-1) + (e^{\beta\phi_s}-\beta\phi_s-1)]^{1/2} \qquad (2.20)$$

where n_0 and p_0 are respectively the equilibrium majority and minority carrier densities in the silicon bulk, and $\beta = q/kT$ is the inverse of the thermal voltage. The space charge Q_{sc} is then obtained from (2.13) and (2.20). It has been shown [8] that half this space charge lies within $\sqrt{2}$ x Debye length of the surface in accumulation, and that the entire accumulation layer width is a few times the Debye length. Impurity redistribution at the silicon surface due to thermal oxidation and ion implantation results in non-uniform doping near the surface. This may cause deviations in the surface potential and the surface electric field. Ancona [10] has suggested that Q_{sc} may be dependent on both the surface potential and the surface electric field. This would affect the dynamic characteristics of the device.

The built-in potential across a p^+-n junction or one-sided abrupt junction is [9]

$$V_{bi} = (1/\beta)\ln(N_A N_D/n_i^2) \qquad (2.21)$$

where N_A is the acceptor density of the p^+-grid, and n_i is the intrinsic carrier density of silicon.

The Si/SiO$_2$ interface-state energy versus surface potential relationship is obtained with reference to Figure 2.5 as follows:

$$E_{ss} = E_F - E_{vs} \qquad \text{(assuming that the interface states are pinned at } E_F\text{)}$$

$$= E_g/2 + \phi_B + \phi_s \qquad \phi_s < 0 \qquad (2.22)$$

where it can easily be shown that

$$\phi_B = (1/\beta)\ln(N_D/n_i) \qquad (2.23)$$

From the charge neutrality condition and Kirchhoff's voltage law, the relationship between the surface potential and the gate voltage is obtained as [6]

$$\phi_s = \int_{V_{FB}}^{V_G} (1-C_{MO}/C_I)\,dV_G \qquad (2.24)$$

where C_{MO} is the measured quasi-static capacitance. The use of V_{FB} to determine this relationship results in a fictitious dependence of energy on temperature, interface-state density, doping profile and oxidation conditions [6], since the energy is calculated using ϕ_s. This problem is avoided in the QSCV technique, where the above relationship is determined empirically.

2.3 Admittance Parameters

The total insulator capacitance is given by

$$C_I = (1/C_{Ox}+1/C_N)^{-1} \tag{2.25}$$

where $C_{Ox} = \varepsilon_{Ox}/T_{Ox}$ is the oxide capacitance, and $C_N = \varepsilon_N/T_N$ is the nitride capacitance.

The quasi-static space charge capacitance C_{sc0} is obtained by invoking the quasi-static approximation:

$$C_{sc0} \cong - dQ_{sc}/d\phi_s \tag{2.26}$$

which gives [9]

$$C_{sc0} = (q\varepsilon_{Si}n_o\beta/2)^{1/2}\left|\frac{(p_o/n_o)(1-e^{-\beta\phi_s}) + (e^{\beta\phi_s}-1)}{[p_o/n_o)(e^{-\beta\phi_s}+\beta\phi_s-1) + (e^{\beta\phi_s}-\beta\phi_s-1)]^{1/2}}\right| \tag{2.27}$$

where $p_on_o = n_i^2$, and $n_o = N_D$. From a consideration of the geometric curves, it can easily be shown that the quantity within the square root in the denominator is greater than or equal to zero for all ϕ_s. The space charge capacitance is sensitive to band-bending fluctuations in accumulation and inversion due to the exponential dependence of the carrier densities on ϕ_s in these regions. The fluctuations themselves are greatest in depletion and decrease towards accumulation and more significantly towards inversion [5,6].

The depletion capacitance of the p$^+$-n junction is [9]

$$C_{pn} = (q\epsilon_{Si}N_D\beta/2)^{1/2}(\beta V_{bi}+\beta\phi_s-2)^{-1/2} \tag{2.28}$$

The Si/SiO$_2$ interface-state admittance is

$$Y_{ss} = j\omega C_p + G_p \tag{2.29}$$

where C_p and G_p are the parallel interface-state capacitance and conductance respectively, and ω is the angular frequency. In Lehovec's Model [11], these are given as

$$C_p = qN_{ss}\tan^{-1}(\omega\tau_n)/(\omega\tau_n) \tag{2.30}$$

$$G_p = qN_{ss}\ln(1+\omega^2\tau_n^2)/(2\tau_n) \tag{2.31}$$

where N_{ss} is the interface-state density; and τ_n is the majority carrier time constant of a continuum of interface states, given by [12]

$$\tau_n = [v_{th}\sigma_n N_D \exp(\ \phi_s)]^{-1} \tag{2.32}$$

where v_{th} is the thermal velocity of electrons, and σ_n is the majority carrier capture cross-section of the interface states. With $\omega \rightarrow 0$, the quasi-static capacitance and conductance are obtained from (2.30) and (2.31) as respectively

$$C_{p0} = qN_{ss} \tag{2.33}$$

$$G_{p0} = 0 \tag{2.34}$$

In deriving the Lehovec's Model, the interface-state charge Q_{ss} has been assumed to be spatially constant. But due to charge non-uniformities, this charge and hence ϕ_s may be randomly distributed. Expressions for C_p and G_p for the case of statistically varying ϕ_s have been derived by Deuling et al. [12]. For the quasi-static and high frequency cases, both models yield the same values for C_p and G_p. Therefore, the statistical model is not of relevance in the QSCV technique and is not shown here. However, it is important for the conductance technique, since it provides an accurate fit for the experimental curves, which are broader than those given by (2.31).

2.4 Conduction Processes

Various charging and discharging mechanisms for different electric field regimes and oxide thicknesses have been proposed in the literature [2,13-18]. These mechanisms for the thin oxide MNOS device are summarized in Figure 2.8. Positive bias charging involves one or more of the following with respect to the oxide: direct band-to-band tunneling (A); modified Fowler-Nordheim tunneling (B), in which the entire oxide and part of the nitride are penetrated; and direct tunneling to traps at the oxide-nitride interface (C). With respect to the nitride, it involves trapping of electrons throughout the nitride which are transported by means of one or more of the following mechanisms: Poole-Frenkel detrapping (field-assisted thermal emission) (D), field or tunnel emission (D) and ohmic current (E). For thick (> 50 Å) oxide devices, Fowler-Nordheim tunneling, that is tunneling of charges through the triangular oxide energy band to the oxide conduction band,

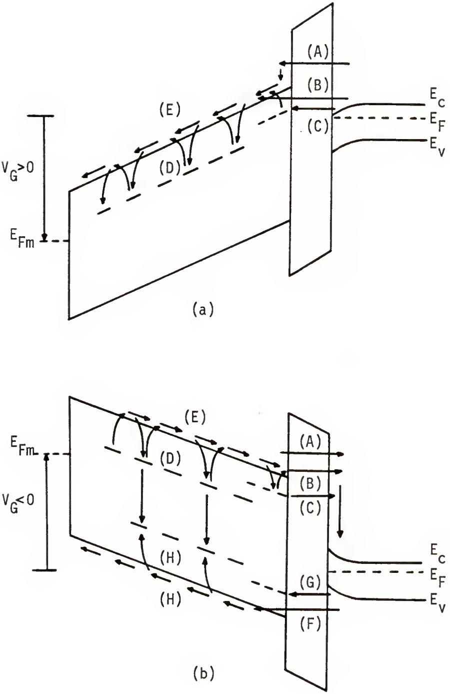


Fig.2.8 Mechanisms for (a) charging and (b) discharging in the MNOS device. (See text for explanations.)

takes place. Negative bias discharging involves the same processes in the opposite direction and, additionally, the following hole conduction processes: modified Fowler-Nordheim tunneling (F) and direct tunneling to traps at the oxide/nitride interface (G) with respect to the oxide and hole transport, trapping and recombination with trapped electrons (H) with respect to the nitride. No hole detrapping is involved in hole conduction.

Oxide/nitride interface traps empty with a time constant less than the tunneling time constant. Therefore, they are referred to as "traps assisting in charge conduction." Ferris-Prabhu [19] reports a maximum tunneling-to-traps distance of 35 Å in the nitride, while Yamamoto et al. [20] report this distance to be 33 Å, 45 Å and 56 Å for oxide thicknesses of 15 Å, 20 Å and 25 Å respectively.

The mechanism for Poole-Frenkel detrapping under an applied electric field is the Poole-Frenkel barrier lowering illustrated in Figure 2.9, where E_t is the energy level of the trap from the nitride conduction band edge; E_{cN} and E'_{cN} are respectively the nitride conduction band edge energy levels before and after the application of the field; and $\Delta\phi_{PF}$ is the potential barrier lowering, given by [9]

$$\Delta\phi_{PF} = \beta_{PF}F_N^{1/2} \tag{2.35}$$

where β_{PF} is the Poole-Frenkel constant, given by

$$\beta_{PF} = [q/(\pi\epsilon_N)]^{1/2} \tag{2.36}$$

With respect to the oxide/nitride interface traps, Poole-Frenkel barrier lowering applies only for positive gate voltage, since electron

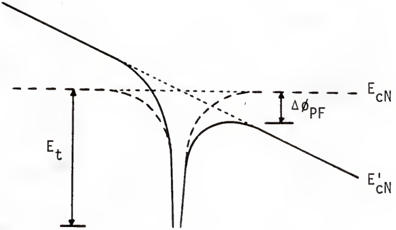


Fig.2.9 Poole-Frenkel barrier lowering in the silicon nitride under an applied field.

transport is into the oxide for negative gate voltage. With respect to the nitride traps, Poole-Frenkel barrier lowering enhances electron movement towards the gate for positive gate voltage and away from it for negative gate voltage. The trap level from which detrapping takes place depends also on the temperature.

The oxide current density is

$$J_{0x} = J_A + J_B + J_C \tag{2.37}$$

where J_A is the Fowler-Nordheim tunneling current density, J_B is the direct tunneling-to-traps current density, and J_C is the direct band-to-band tunneling current density. The Fowler-Nordheim current density is [2,9]

$$J_A \propto (F_{0x}^2/\phi_{Ba})\exp(-A_1\phi_{Ba}^{3/2}/F_{0x}) \tag{2.38}$$

where A_1 is a constant; and $\phi_{Ba} = \phi_{0xS}$ for electrons during charging, or $\phi_{Ba} = \phi_{0xN}$ for electrons during discharging, or ϕ_{Ba} = the Si-SiO$_2$ valence band barrier for holes during discharging. The Si-SiO$_2$ conduction band barrier for electrons is modified by the image force effect. The reduction in the barrier height is given by [2,9]

$$\Delta\phi_{Ba} = [qF_{0x}/(4\pi\varepsilon_{0x})]^{1/2} \tag{2.39}$$

The direct tunneling-to-traps current density is [21]

$$J_B \propto N_t\exp(-B_1/|F_{0x}|) \tag{2.40}$$

where N_t is the nitride trap density, and B_1 is a constant. The direct band-to-band tunneling current density is [22]

$$J_C \propto (1/\beta)\exp[-C_1T_{0x} - \beta(\phi_{0xS}-V_{0x})](e^{\beta\phi_{0xN}-1}) \qquad (2.41)$$

where C_1 is a constant.

The nitride current density is

$$J_N = J_1 + J_2 + J_3 \,. \qquad (2.42)$$

where J_1 is the Poole-Frenkel current density, which dominates at high temperatures and high fields; J_2 is the field (or tunnel) emission current density, which dominates at low temperatures and high fields and is independent of the temperature; and J_3 is the ohmic current density due to thermally excited electrons hopping from one isolated state to the next, which contributes to J_N at low fields and moderate to high temperatures. The Poole-Frenkel current density is [2,9]

$$J_1 \propto F_N\exp[-\beta(E_t-\Delta\phi_{PF})] \qquad (2.43)$$

where E_t is in eV. The field emission current density is [9]

$$J_2 \propto F_N^2\exp(-F_{0x}/F_N) \qquad (2.44)$$

The ohmic current density is [23]

$$J_3 \propto F_N\exp(-\beta\Delta E_{ae}) \qquad (2.45)$$

where ΔE_{ae} is the activation energy of electrons in eV.

The dominant oxide current is direct tunneling current in thin (< 50 Å) oxide devices and Fowler-Nordheim tunneling current in thick oxide devices [14,20,24]. In thin oxide devices, direct tunneling may take place even during the application of low Read voltages [15,16]. Maes and Van Overstraeten [21] suggest that the oxide field decreases with time of the applied pulse due to the injected charge opposing the field, which results in a transition from Fowler-Nordheim tunneling to trap assisted tunneling for thin oxide devices. Trap assisted charge injection is significant only at low fields, because shallow traps, which are instrumental for this mechanism, will be aligned with the silicon conduction (or valence) band only at low fields. Chang [22] suggests that the switching process is dominated initially by direct band-to-band tunneling and then by modified Fowler-Nordheim tunneling. The avalanche breakdown field for silicon is $3\text{x}10^5$ V/cm. Then, from (2.2), the oxide field works out to about 10^6 V/cm. Charge conduction in the oxide is negligible at this field strength even at high temperatures. However, mobile ions can transport through the oxide at high fields, giving rise to electrical instability. Also, the dielectric relaxation time for a good quality thermal oxide is about four days [6], which implies that the silicon dioxide cannot be at equilibrium with any applied a.c. field. Therefore, the oxide current is essentially a tunneling rather than a conduction current.

Yamamoto et al. [20] report that the discharge process is exponential with time until saturation for small (< 50 Å) tunneling distances. For larger tunneling distances, the discharging is exponential at first and then logarithmic until saturation, the transition occurring at a sub-second value. The larger the tunneling distance, the sooner the transition takes place.

For a positive charging pulse, electron injection from the silicon is more probable than hole injection from the gate, considering the thin oxide layer and the barrier potential between the nitride valence band edge and the metal Fermi level. Carrier injection from the gate is negligible for aluminum gates but may not be for some other metallic gates [25]. Al-gate MNOS devices with 500 Å oxides have shown a very small amount of electron injection from the gate but no hole injection, while n-doped poly-silicon gate MNOS devices with 500 Å oxides have shown hole injection from the gate but no electron injection [26].

Holes have higher mobility in the nitride than electrons due to hole traps being shallower than electron traps [25,27]. Hence, higher nitride currents are obtained for negative bias than for positive bias. In silicon dioxide, electron traps are much deeper than in the nitride, so that communication with them is inhibited [27]. Also, the electron trap density in silicon dioxide is five orders of magnitude lower than in the nitride (i.e. the silicon dioxide is more defect free), and the electron effective mass is very low, while holes form self-trapped polarons [28]. Hence, electron transport dominates in the oxide.

The greater the magnitude or width of the Write or Erase pulse, the deeper is the penetration of carriers into the nitride [29]. However, beyond a certain limit, conduction through the device begins to take place. Also, the stored charge can get reduced due to the injection of the opposite kind of charge from the gate. The thinner the oxide and the nitride layers, the smaller are the magnitude and width of Write/Erase pulses required. A thinner oxide layer also implies a higher oxide current and more stored charge. But if the oxide layer is too thin (say < 15 Å), retention is degraded, and the Si/SiO_2 interface states may communicate with the nitride.

CHAPTER III
MNOS ADMITTANCE AND MECHANICS OF OPERATION

3.1 Derivation of the Small Signal Equivalent Circuit

The cross-section of the p^+-gridded MNOS capacitor is shown in Figure 3.1 (a). When no memory charge is present, the charge neutrality condition with respect to the n-region gives

$$Q_{M1} = - Q_{ss} - Q_{sc} - Q_{pn} \tag{3.1}$$

where Q_{M1} is the charge in the metal gate due to the n-region charges Q_{ss}, Q_{sc} and Q_{pn} (the p^+-n junction charge). The small signal current flowing out of the gate due to the charge in the n-region is obtained by differentiating (3.1); thus,

$$\begin{aligned} i_{G1} &= - dQ_{M1}/dt = dQ_{ss}/dt + dQ_{sc}/dt + dQ_{pn}/dt \\ &= - i_{ss} - i_{sc} - i_{pn} \end{aligned} \tag{3.2}$$

where t is the time; and i_{ss}, i_{sc} and i_{pn} are the small signal currents flowing into the silicon due to the corresponding charges. We can write (3.2) as

$$i_{G1} = - Y_{ss}\phi_s' - j\omega C_{sc}\phi_s' - j\omega C_{pn}\phi_s' \tag{3.3}$$

where C_{sc} is the space charge capacitance, and ϕ_s' is the a.c. component of the surface potential.

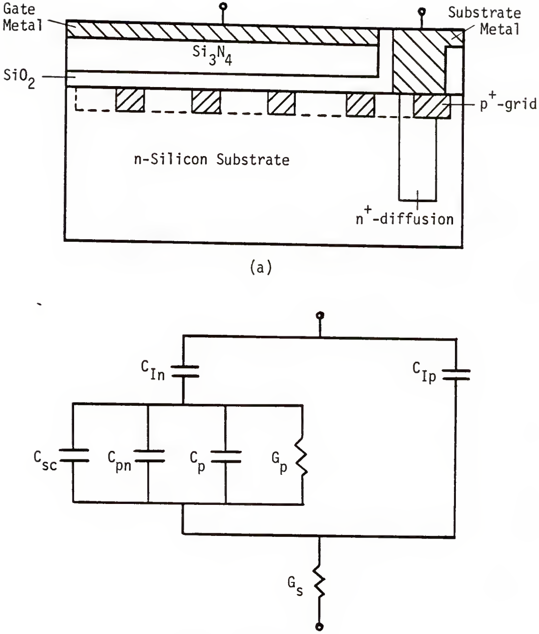


Fig.3.1 The p^+-gridded MNOS capacitor: (a) the cross-section and (b) the small signal equivalent circuit.

We can write the total surface potential ϕ_s^*, the total gate voltage V_G^* and the total voltage across the series resistance G_s, V_{Gs}^*, as summations of their respective d.c. and a.c. components; thus,

$$\phi_s^* = \phi_s + \phi_s' \tag{3.4}$$

$$V_G^* = V_G + V_G' \tag{3.5}$$

$$V_{Gs}^* = V_{Gs} + V_{Gs}' \tag{3.6}$$

where $\phi_s' = \phi_{sm}e^{j\omega t}$, $V_G' = V_{Gm}e^{j\omega t}$ and $V_{Gs}' = V_{Gsm}e^{j\omega t}$ are the a.c. components; and ϕ_{sm}, V_{Gm} and V_{Gsm} are their respective amplitudes. From Kirchhoff's voltage law,

$$V_G^* = Q_{M1}/C_{In} + \phi_{MS} + \phi_s^* + V_{Gs}^* \tag{3.7}$$

where C_{In} is the insulator capacitance between the n-region and the gate. Differentiating (3.7) with respect to time and substituting from the time differentials of (3.4) - (3.6), we obtain

$$j\omega V_G' = - i_{G1}/C_{In} + j\omega\phi_s' + j\omega V_{Gs}' \tag{3.8}$$

Substituting from (3.3), we obtain

$$V_G' = - (1/j\omega C_{In} + 1/Y_1)i_{G1} - i_G/G_s \tag{3.9}$$

where $Y_1 = Y_{ss} + j\omega C_{sc} + j\omega C_{pn}$ (i.e. the parallel combination), and $i_G = - V_{Gs}'G_s$ is the total small signal current flowing out of the gate. Hence, writing Y as the admittance of Y_1 and C_{In} in series, we obtain

$$V_G' = - i_{G1}/Y - i_G/G_s \tag{3.10}$$

Since the total charge on the gate is the algebraic sum of Q_{M1} and the charge on the gate due to the insulator capacitance C_{Ip} between the p^+-grid and the gate, we can consider C_{Ip} to be providing a current path to the substrate parallel to that provided by Y. Hence, equating potentials,

$$i_{G1}/Y = i_G/(Y+j\omega C_{Ip}) \tag{3.11}$$

Substituting in (3.10), we obtain

$$V_G' = -\ [1/(Y+j\omega C_{Ip}) + 1/G_s]i_G \tag{3.12}$$

Therefore, the equivalent circuit comprises of G_s in series with the parallel combination of C_{Ip} and Y and is shown in Figure 3.1 (b).

3.2 Equivalent Circuit with Reference to the Energy Band Diagram

The equivalent circuit with reference to the energy band diagram is shown in Figure 3.2, where C_D, C_{INV}, G_{grs}, G_{gr} and G_d are respectively the depletion and inversion capacitances, the conductances associated with the depicted generation-recombination (G-R) processes at the Si/SiO$_2$ interface states and the silicon bulk traps and the minority carrier diffusion and drift conductance, which is negligible at room temperature; E_T is the silicon bulk trap energy level; and C_p' and G_p' are respectively the interface-state capacitance and conductance for minority carrier capture and emission, which are usually negligible compared to C_{INV}. During inversion, the silicon surface becomes continuous with the p^+-grid, which is connected to the n-substrate, effectively shorting the p^+-n junction capacitance.

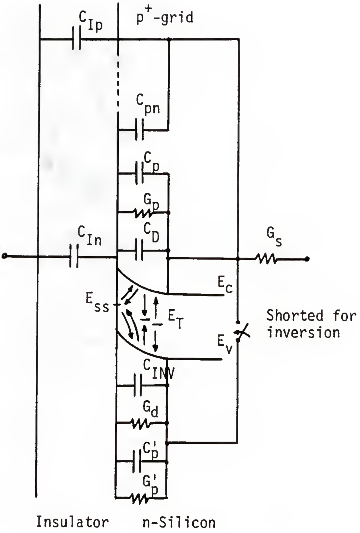


Fig.3.2 Equivalent circuit of the p^+-gridded MNOS capacitor with reference to the energy band diagram.

The series resistance is due to the gate and substrate contacts, the n-silicon bulk and built-in potentials resulting from non-uniform doping. The higher the donor density, the lower is the series resistance. A 250 Ohm series resistance is considered normal [30]. The series resistance is negligible at low frequencies, and its effect at high frequencies can be reduced by careful fabrication and making corrections to measurements. On the other hand, some series resistance is desirable, as it would reduce the voltage across the oxide, thus improving the endurance of the device.

At a given gate voltage or surface potential value, the interface-state capacitance arises from the capture and emission of carriers from that interface trap level which is intersected by the Fermi level. The space charge capacitance, which is equal to C_D in depletion and C_D+C_{INV} in inversion, arises from the charge flow (in and out of the space charge layer) needed to maintain thermal equilibrium at every point in the semiconductor during an applied a.c. signal. The interface-state conductance arises from the phase lag between the interface-state response and the applied a.c. signal. In one half of the a.c. cycle, the traps are too slow to be filled, so that some traps below E_F are empty, and in the other, some traps above E_F are too slow to empty. Both processes result in a loss of energy. This energy is supplied by the a.c. source, which as phonons heats up the silicon lattice. The loss is represented by the interface-state conductance G_p. The conductances G_d, G_{gr} and G_{grs} also arise from the phase lag between the respective processes and the a.c. signal.

3.3 Carrier Response

Carrier response time is the time required for carriers to be supplied to or removed from the silicon space charge layer in response to an applied disturbance. Carriers will respond to an a.c. signal if the angular frequency is very much less than the inverse of the response time. Carrier lifetime is the time required for the carriers to return to thermal equilibrium after a disturbance by changing their concentrations through G-R. For minority carriers, response time depends on lifetime. Traps alternately emit holes and electrons. Therefore, free carrier lifetime τ_0 depends on both majority and minority carrier processes. If majority carrier processes take longer than minority carrier processes, then τ_0 will be determinied by the former, and vice-versa. Free carrier lifetimes of 4 - 5 μs in the bulk [31,32] and 1 μs at the surface [6,8] and a minority carrier response time between .01 and 1 sec at room temperature [6] have been reported.

There are four mechanisms for minority carrier transitions to and from the inversion layer for the p^+-gridded device. They are

(i) Bulk trap G-R near the midgap in the silicon quasi-neutral region close to the back contact, diffusion through this region and drift through the depletion layer. This is significant only at high temperatures.

(ii) Bulk trap G-R in the depletion layer near the midgap when E_F intersects E_T. The centers are efficient if $E_T \approx E_i$. Then, the minority carrier lifetime ≈ the majority carrier lifetime.

(iii) Surface G-R in the interface traps. This is significant only in the midgap to weak inversion bias range when the interface-state density is greater than 10^{10} $eV^{-1}cm^{-2}$ [6].

(iv) From the p^+-grid across the p^+-n junction. This is the dominant source.

It should be noted that the silicon back contact, though ohmic for majority carriers (i.e. connected to the silicon conduction band), is blocking for minority carriers. Therefore, no minority carriers are injected from the external circuit.

3.4 C-V and G-V Characteristics

The quasi-static C-V curves reported in the literature for both MOS and MNOS manifest a valley in the depletion to weak inversion region and saturate to the insulator capacitance value towards strong accumulation and strong inversion. The high frequency C-V curves saturate to the insulator capacitance value in strong accumulation and to a minimum value in strong inversion. This high frequency behavior is due to the majority carriers being able to respond to the high frequency (about 1 MHz) signal and the d.c. bias signal and the minority carriers being able to respond only to the latter. Hence, in inversion, the capacitance remains at its minimum depleted value. The C-V curve decreases with frequency significantly in inversion and barely in accumulation, although in some reported results, a discernible decrease in accumulation also can be seen [5,33]. The inversion side behavior has been attributed to G-R loss [6] and frequency dependence of the minority carrier concentration [34], while the slight decrease in the accumulation side has been attributed to the series resistance [6].

The metal-semiconductor work function difference ϕ_{MS}, the fixed oxide charge and the stored nitride charge cause a C-V shift without distortion from the ideal along the voltage axis. For ϕ_{MS}, the shift direction has the same sign as ϕ_{MS}. For each of the charges, the shift direction is opposite in sign to the charge. Slow interface states and any mobile oxide charge will follow the d.c. bias but not the a.c. signal. The mobile charge gives rise to hysterisis of the C-V curves, while the slow states cause stretch-out (distortion) of the high frequency C-V curve compared to the quasi-static C-V curve. In both cases, the minimum measured capacitance C_{min} is unaffected. Fast interface states can follow both the d.c. bias and the a.c. signal (up to a certain frequency). These cause a C-V shift, a high frequency C-V stretch-out and a change in the C_{min} value. The C-V shift due to charges was dealt with in Chapter II with respect to flatband and threshold voltages. The C-V stretch-out and the change in C_{min} value can be explained as follows: At a given bias, fewer interface states respond to the high frequency signal than to the low frequency signal. This effect is more significant in accumulation and inversion than in depletion. Hence, a broader and lower C-V curve is obtained.

Impurity redistribution also affects C-V curves [6,35]. For example, oxidation results in the boron (acceptor) density being reduced near the surface. This causes the depletion width to be larger and hence the capacitance to be smaller for a given bias. Therefore, both the C-V curve in depletion and C_{min} will be lower. Also, inversion occurs sooner, with the result that the C-V curve in inversion will be higher. The net effect is a lower C-V curve, shifted in the negative voltage direction. On the other hand, oxidation results in the phosphorous (donor) density being increased near the surface. This gives a higher C-V curve (smaller depletion width), shifted in the negative voltage direction (delayed inversion).

The G-V curves reported in the literature increase with frequency from a negligible quasi-static conductance value [5,30,32,33]. It is constant in strong accumulation and strong inversion and is peaked in the depletion to weak inversion region. Singh and Srivastava [33] report a dented conductance peak for MOS devices. The G-V curve at a given bias range is characterized by the dominant loss in that range. Any significant capacitance variations may also be reflected in this curve.

In strong accumulation and strong inversion, the space charge is very close to the silicon surface. That is, its width is small. Therefore, the space charge capacitance C_{sc} is high. The interface states are completely filled or completely empty, and no communication with them takes place. Therfore, they experience no loss and contribute no capacitive value. Hence, C_{sc} dominates in these regions. At low frequencies, the effect of the series resistance is negligible. Therefore, the device capacitance approaches the insulator capacitance in value. In depletion, the space charge capacitance decreases as depletion width increases, and in inversion, it increases again due to the increase of the inversion charge. The interface-state capacitance makes a significant contribution to the device admittance in depletion and weak inversion. This contribution decreases with frequency and becomes negligible at high frequencies.

At intermediate frequencies, several losses come into play. The interface-state conductance G_p versus the gate voltage is sharply peaked, trailing off to zero on either side [6]. In accumulation, there is an ample supply of majority carriers, so that no loss is experienced. In depletion, the majority carrier density is low, resulting in an energy loss. In inversion, the majority carrier density is so low, that no capture or emission takes place, and again the loss is zero. In the weak inversion region, within about 3kT/q of E_i, G-R through bulk traps as E_F intersects E_T and through interface traps as E_F intersects the interface states near the midgap takes place. This region of bias is therefore too complex for analysis and must be avoided. In strong inversion, G-R through bulk traps dominates. The surface G-R conductance G_{grs} versus the gate voltage is peaked, but the bulk G-R conductance G_{gr} is not, since it is independent of the gate voltage [6].

It is worthwhile mentioning the insulator polarization effect on C-V curves reported by Tutto and Balazs [36]. The Si-H and N-H bonds in silicon nitride are inhomogeneously distributed, which gives an inhomogeneous built-in field. This field provides macroscopic potential valleys as preferred places for quasi-free charge carriers (of the order of 10^{11} cm^{-2}) in the nitride, giving rise to charge motion along most of the length of the silicon nitride. This effect, which is referred to as "Insulator Polarization," gives higher C_{MO} values, a greater slope to the high frequency C-V curve in depletion and an order of magnitude higher interface-state density values, especially towards accumulation.

3.5 G-ω Characteristics

For a given surface potential value, G_p/ω versus log ω peaks at a frequency given by [6]

$$\omega_m = 1.98/\tau_n \qquad (3.13)$$

Considering band-bending fluctuations, this frequency is [5,6]

$$\omega_m = 2.5/\tau_n \qquad (3.14)$$

For $\omega << \omega_m$, the interface states respond to the a.c. signal immediately, and therefore the loss is zero. For $\omega >> \omega_m$, the interface states do not respond to the a.c. signal at all. Again, the loss is zero. When the frequency approaches ω_m, the loss increases due to the increase in the phase lag between the applied a.c. signal and the interface-state response. The experimental G_p/ω versus log ω curves are broader than those predicted by (2.31). The reasons for this are

(i) Statistical fluctuation in the surface potential.

(ii) Fluctuations in the majority carrier capture cross-section.

(iii) The spatial isolation of interface traps, even though they are in an energy continuum. For a trap density of 10^{10} to 10^{12} per sq.cm., the seperation of adjacent traps is 10^3 to 10^2 Å [6]. This results in each interface trap energy level having an individual peaked conductance-frequency response, as opposed to a composite frequency response when the traps at a given level are able to communicate with one another. The cumulative effect is a broadening of the conductance peak.

The G_p/ω versus log ω curve at a given surface potential value is asymmetrical about the frequency ω_m, the curve being higher on the high frequency side [6]. This is attributable to the spatial isolation of traps, with those at higher energy levels undergoing greater loss and having higher peak frequencies than those at lower energy levels.

A loss called the capture and emission loss of bulk traps, similar to G_p for interface traps, becomes more significant than G_{gr} at a very high frequency [6]. But, for low bulk trap densities, this frequency is beyond the frequency range of practical interest. Chan et al. [32] have reported that G_{gr} versus $1/\omega$ increases linearly and saturates at high $1/\omega$ values.

3.6 Energy Level Behavior

As the bias voltage V_G is changed, E_{cs}, E_{vs} and E_{ss} change with respect to E_F. At each V_G (or ϕ_s) value, the measured values of the device capacitance and conductance reflect those corresponding to the trap level intersected by E_F. Interface states above E_F will be empty, and those below E_F will be filled. Interface traps empty by emitting majority carriers (for n-type), and vice-versa. In inversion, both majority carrier emission and minority carrier capture (i.e. G-R) may be involved in emptying of traps.

CHAPTER IV
DEVICE FABRICATION, EXPERIMENTS AND ANALYSIS

4.1 Device Fabrication and Features [37]

The p^+-gridded MNOS devices used in this research were fabricated at the Harris Semiconductor Division and had a 3-5 Ohm-cm <100> orientation n-silicon material with a donor density of 2×10^{15} cm^{-3} for the substrate. An oxide was grown on this substrate in steam at 1100^oC to a thickness of 5000 Å. The oxide was etched, and an n^+ diffusion was carried out to provide a contact to the substrate. A grid pattern was also etched into the oxide, and a boron diffusion was carried out by the standard thermal diffusion procedure to obtain a 1.2×10^{-3} ohm-cm p^+-grid with an acceptor density of 1×10^{20} cm^{-3}. The oxide was then regrown using a 900^oC 90 minute steam process. The grid may also be produced by post deposition ion implantation [4]. A mask was used to pattern and etch the oxide to the required thickness. The nitride was then deposited using the AMT MODEL 5100 Low Pressure Chemical Vapor Deposition (LPCVD) system. The deposition parameters used were a temperature of 756^oC, a pressure of .438 torr and a deposition rate of 25.9 Å/min. After making contact cuts to the n^+- and p^+-regions, the entire wafer was metallized by a standard aluminum deposition. The aluminum was then patterned and etched by a mask which defined the MNOS capacitor, the probe pads and the connections. The processing variations used were two pre-nitride NH_4OH dips, two different ammonia/dichlorosilane (DCS) ratios and a post-nitridation H_2 anneal. These are summarized in Table 4.1.

A nitride thickness uniformity of better than 1.0 % for a wafer and 2.8 % overall was achieved. The oxide and nitride thicknesses were 20 Å and 465 Å respectively. The n-silicon and the p^+-grid had areas of 5.450×10^{-4} and 2.078×10^{-4} cm^2 respectively. The p^+-grid was 1 µm deep and electrically shorted to the substrate.

4.2 Measurements

The C-V measurements were carried out using an HP 4140B pico-Ammeter, an HP 4274A LCR meter, an HP 4275A LCR meter and an HP 4191A impedance analyzer. Write/Erase cycling was done on the Harris MNOS capacitor exercisor. Heating for temperature measurements was done in a Stathem oven. Bias stress was provided from an HP 6112A d.c. power supply.

For Write/Erase cycling, ∓26 V 10 µs pulses were used with a 1 µs pause between pulses. Programming was done using ∓20 V 10 s saturating pulses. Several bipolar preconditioning pulses were applied before each retention measurement to ensure that the programming time was the same for all measurements. In the retention measurements, the gate voltage V_{TH}' corresponding to a fixed capacitance value was taken as a measure of the threshold voltage. The Read disturb effect was minimized by keeping the voltage sweep time to a minimum.

Table 4.1 Processing Variations Used in the Fabrication of the MNOS Devices

Process Step	Device #			
	3	4	7	15
Two 10 sec pre-nitride NH_4OH dips	x	x	x	-
9:1 ammonia:DCS ratio	x	-	x	x
100:1 ammonia:DCS ratio	-	x	-	-
$900^{o}C$ 30 min H_2 anneal	-	-	x	-

4.3 QSCV Analysis and Results

Figure 4.1 shows the C-V curves for device #3, where pre-nitride NH_4OH dips and a low ammonia/DCS ratio have been employed. The C-V curves for the other devices are similar. The coincidence of the 200 kHz and 400 kHz curves in the depletion region and of the 400 kHz and 1 MHz curves in the weak inversion region suggests that, at 400 kHz, the parameters which contribute frequency dependent components to the measured capacitance C_M have a negligible effect on it in these bias regions. Thus, the 400 kHz curve, whose only significant components in the depletion to early weak inversion bias range are C_{sc}, C_{pn}, C_{In} and C_{Ip}, can be used as the high frequency curve for the QSCV analysis.

From the foregoing and Figure 3.1 (b), we can write

$$C_M = [1/C_{In} + 1/(C_{sc}+C_{pn})]^{-1} + C_{Ip} \qquad (4.1)$$

The C_{In} and C_{Ip} values are obtained from the measured insulator capacitance value based on their aerial ratios. A FORTRAN program (see Appendix B) was written to compute the minimum experimental value of $C_{sc}+C_{pn}$ from (4.1) and the minimum value of C_M and to obtain the donor density and built-in potential values by matching this minimum value with the theoretical minimum value given by (2.27) and (2.28) using a recursive technique. The experimental $C_{sc}+C_{pn}$ values at different gate voltage values, which are obtained by substituting the C_M values in (4.1), are compared with the corresponding theoretical values to determine the surface potential versus gate voltage relationship. For

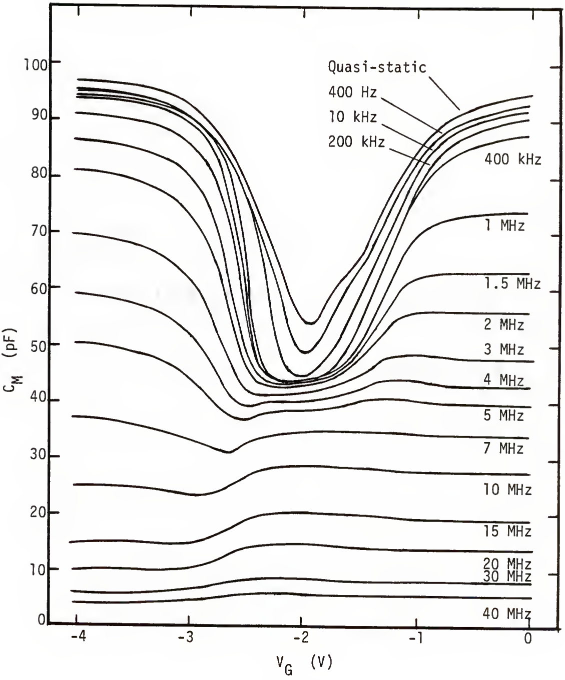


Fig.4.1 C-V curves for the p^+-gridded n-silicon MNOS capacitor with frequency as parameter.

the quasi-static case, all losses are zero, and the effect of G_s is negligible. Therefore, from Figure 3.1 (b), we can write

$$C_{MO} = [1/C_{In} + 1/(C_{sc}+C_{pn}+C_{pO})]^{-1} + C_{Ip} \qquad (4.2)$$

Hence, C_{pO} can be obtained. Then, substituting the value of C_{pO} in (2.33), the value of N_{ss} is obtained for each surface potential value. The energy versus surface potential relationship is obtained from (2.22). Hence, the interface-state density versus energy relationship can be found. Computed parameters for device #3 are shown in Figures 4.2 - 4.5 and below. Similar results were obtained for the other devices.

Donor density = $2.023\text{x}10^{15}$ cm^{-3}
Built-in potential = .891 V
Bulk potential = .306 V

The hump observed in the interface-state density curves is also manifest in the quasi-static C-V curves. Quasi-static C-V curves for MNOS and MOS capacitors fabricated on the same wafer are shown in Figure 4.6. The absence of the hump in the MOS C-V curve suggests that it is associated with the nitride deposition process. The hump can be attributed to the penetration of the thin oxide layer by nitrogen and other impurity atoms, causing changes in the chemical composition at the Si/SiO_2 interface, and/or to structural changes such as bond angle distortions at the Si/SiO_2 interface caused by the nitride deposition conditions.

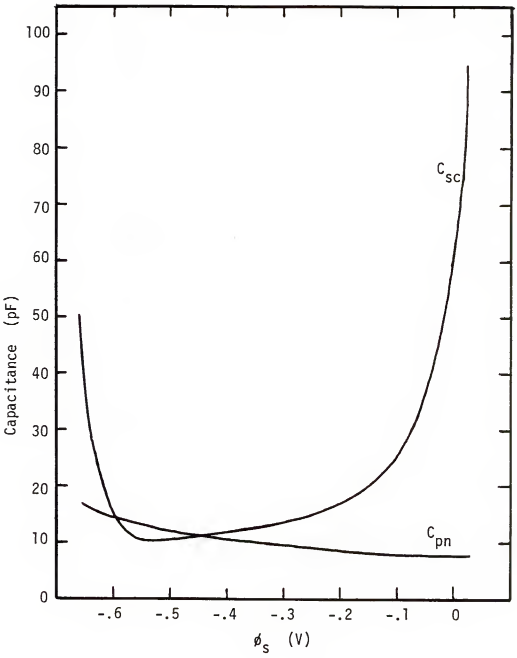


Fig.4.2 Variation of space charge and p^+-n junction capacitances with surface potential for the p^+-gridded n-silicon MNOS capacitor (device #3).

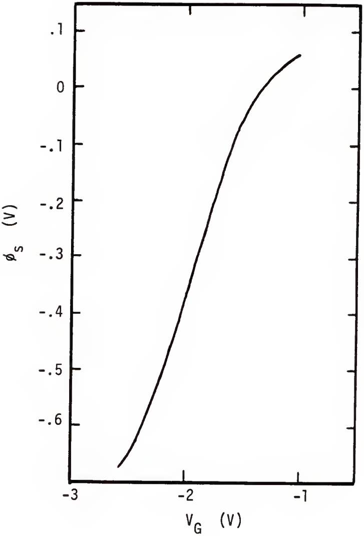


Fig.4.3 Variation of surface potential with gate voltage for the p^+-gridded n-silicon MNOS capacitor (device #3).

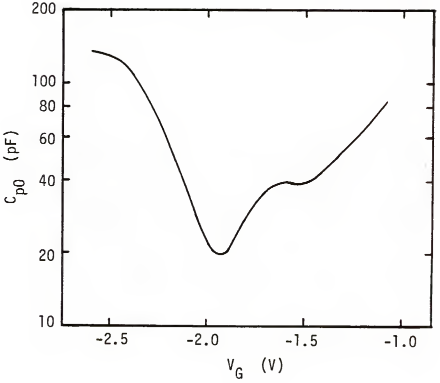


Fig.4.4 Variation of quasi-static interface-state capacitance with gate voltage for the p^+-gridded n-silicon MNOS capacitor (device #3).

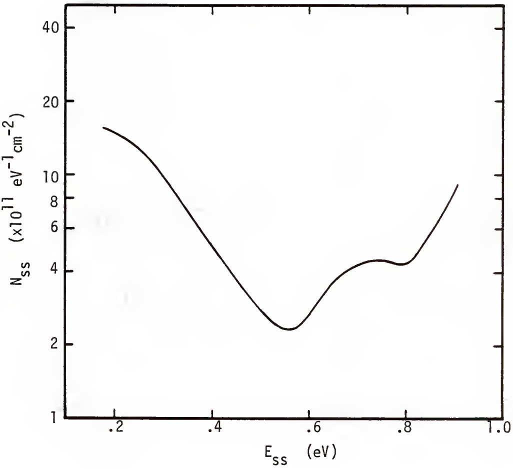


Fig.4.5 The Si/SiO$_2$ interface-state density versus energy curve for the p$^+$-gridded n-silicon MNOS capacitor (device #3).

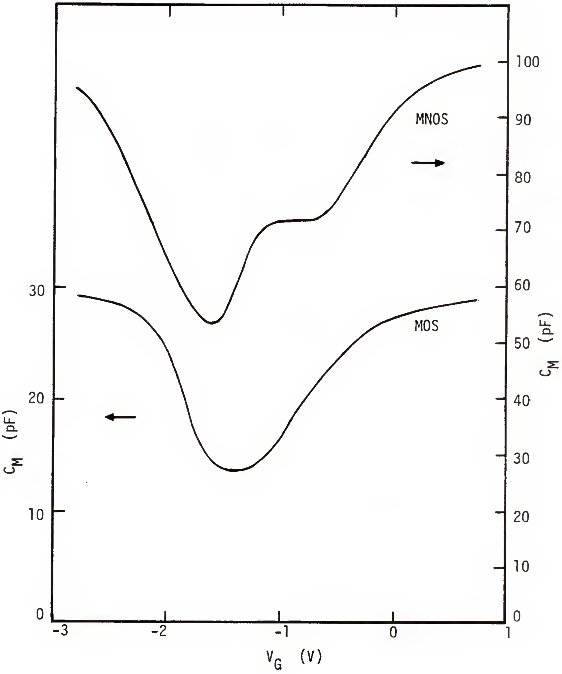


Fig.4.6 Measured C-V curves of MNOS and MOS capacitors fabricated on the same wafer.

4.4 Model Verification

The basic model for the p^+-gridded MNOS capacitor is adequate to accurately simulate the experimental C-V curves up to about 400 kHz. However, it leads to deviations in the high frequency C-V curves in the strong accumulation and strong inversion regions. As quantitative descriptions for G-R conductances do not exist, simulation of the high frequency C-V curves in the inversion region is not possible and is therefore avoided. The deviation in the high frequency C-V curves in the accumulation region is clearly due to an energy loss, which can be represented by a conductance. The C-V curves for several different models were calculated using a FORTRAN program with the appropriate admittance equations for each model. The model whose C-V curves fit the experimental curves best is shown in Figure 4.7, where the conductance G represents the space charge energy loss that causes the deviation in the C-V curves. The other models used involved simple translations of G to other locations in the basic model. These locations were

(i) Between A and B, representing a leakage across the insulator.

(ii) Between B and D, representing a loss parallel to the n-silison and p^+-n junction capacitances.

(iii) Between D and E, representing a loss in the n-silicon substrate.

In Figure 4.7, the admittance between A and E, excluding C_{Ip}, is

$$Y = \frac{1}{1/[1/(1/G + 1/j\omega C_{sc}) + G_p + j\omega(C_p+C_{pn})] + 1/j\omega C_{In}} \qquad (4.3)$$

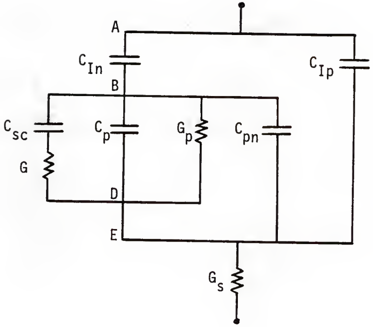


Fig.4.7 Complete equivalent circuit for the simulation of the experimental C-V curves of the p^+-gridded n-silicon MNOS capacitor.

The total device admittance is

$$Y_T = \frac{1}{1/G_s + 1/(Y+j\omega C_{Ip})} \tag{4.4}$$

A FORTRAN program (see Appendix C) was written to calculate the admittance and physical parameters using (2.21), (2.22), (2.23), (2.27), (2.28), (2.30), (2.31) and (2.32) and the device admittance using (4.3) and (4.4). G was found from curve fitting to be frequency dependent. The variation of G with frequency is shown in Figure 4.8. This conductance may be due to a combination of losses, including a loss due to the presence of the p^+-grid. Other data used in the model simulation were the surface potential versus gate voltage curve of Figure 4.3, the interface-state density versus energy curve of Figure 4.5 and the majority carrier capture cross-section versus energy data from [5]. The C-V curves in the depletion to inversion transition region were found to be determined by the value of the series resistance and to be independent of G. Hence, the series resistance could be obtained by curve fitting and was found to be 250 Ohms. The simulated C-V curves are shown in Figure 4.9. It can be seen that these curves compare well with the experimental curves of Figure 4.1.

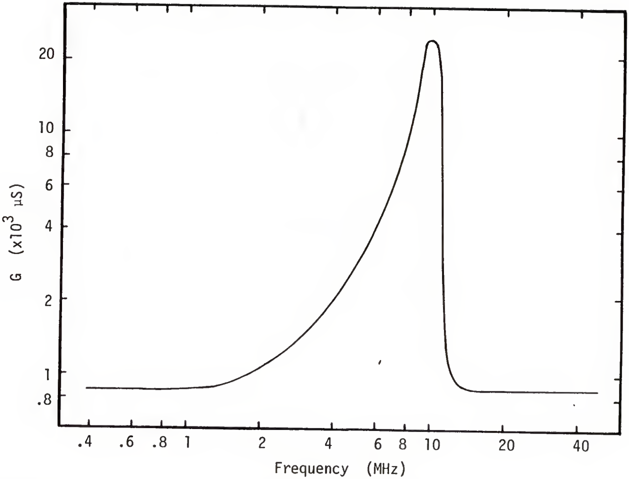


Fig.4.8 Variation of space charge loss conductance with frequency for the p^+-gridded n-silicon MNOS capacitor.

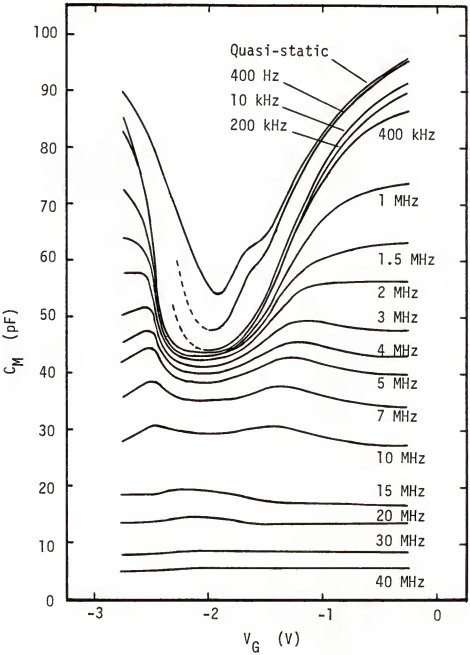


Fig.4.9 Simulated C-V curves of the p^+-gridded n-silicon MNOS capacitor.

CHAPTER V
TRAPS, CHARGES AND THEIR ORIGINS

5.1 Chemical Composition and Structure

5.1.1 Measuring Techniques

The measuring techniques that are being used to study the chemical composition and structure of various regions of the MNOS device are listed below. X-ray Photoelectron Spectroscopy (XPS) [38] and Silicon 2p Photoelectron Spectroscopy (Si 2p PS) [39] are used to analyze the oxide. The Si 2p PS gives better resolution than the XPS. Ellipsometry [40] yields the thicknesses and the refractive indeces of the insulator layers. High Resolution Electron Microscopy (HREM) [41] and Reflection High-Energy Electron Diffraction (RHEED) [42] are used for structural measurements of the oxide. Electron Spin Resonance (ESR) [40] measures paramagnetic defects and impurities in solids. Dangling bonds are paramagnetic due to the single lone-pair electron. Rutherford Backscattering Spectrometry (RBS) [40,43,44] gives the absolute atomic composition versus depth. Therefore, it is used to obtain the Si/N ratio. Nuclear Reaction Analysis (NRA) [43,44] gives the absolute concentration of H and O versus depth. Sputter Auger Electron Spectroscopy (SAES) [43-45] gives better depth resolution than NRA or RBS, but the atomic composition is not absolute. Also, sputtering causes surface roughness, giving a broadening effect to measured interfaces. Infra-Red Absorption (IRA) measurements [43,44] yield light element (e.g. H) bonding configurations in transparent solids (e.g. Si-H and N-H in silicon nitride). Multiple Internal Reflection (MIR) [44] yeilds a more sensitive measurement of Si-H and N-H than IRA.

5.1.2 Silicon Nitride

The structure of amorphous silicon nitride is described [44,46] as being similar to crystalline Si_3N_4 in that the Si atom is coordinated by four N atoms in a distorted tetrahedral arrangement, and the N atom is coordinated by three Si atoms in a nearly planar trigonal arrangement. They differ in that the nearest neighbor distances in the former have a random distribution (i.e. a random distribution of bond angles) about their crystalline counterparts. However, the stoichiometry varies widely from device to device, depending on the deposition conditions, especially the ammonia/DCS ratio.

The composition and structure of the silicon nitride have been studied using the ESR, MIR, NRA, RBS and SAES techniques and energy level calculations [28]. Up to a few atomic percentage (at.%) of O (from contamination of ammonia and DCS gases, from small leaks and from occluded O in the deposition system), up to 10 at.% of H bonded to Si or N (from dissociation of ammonia and DCS gases) and up to 20% excess Si have been reported [28,43,44,47-50]. LPCVD produces near stoichiometric silicon nitride in the bulk. However, the top surface contains some O, giving rise to an oxynitride layer, which extends in a low concentration tail to about 60 Å in the nitride [43,44,49-51]. For silicon nitride deposited at 750-900^{o}C, a 5-30 at.% of N and a

0.5-5 at.% of Si bonded to H atoms, which give rise to an SiN_xH_y composition, have been reported [52]. The subscripts x and y depend on the processing conditions and can be altered by post deposition treatments. The N/Si ratio decreases with depth from the nitride surface, causing the nitride to become silicon rich towards the oxide/nitride interface [53]. The O concentration is uniform in the mid-nitride region and increases towards both ends, while the Si concentration is uniform almost throughout the nitride [54]. The total $\equiv$Si center density in the nitride is 10^{19} cm^{-3} [28].

5.1.3 Silicon Dioxide and Si/SiO_2 and Oxide/Nitride Interfaces

The chemical composition and structure of the oxide and the adjacent interfacial regions have been studied using the ellipsometry, ESR, HREM, IRA, RBS, RHEED, SAES, Si 2p PS and XPS techniques and by comparing electrical measurement data with theoretical models [30,55]. Several different chemical models for these regions have been proposed in the literature.

The oxide/nitride interface and part of the oxide close to this interface (or the entire oxide in the case of thin oxides) are considered to be constituted of an oxynitride of the form SiO_xN_y, where the respective values of x and y change from 0. and 1.33 in the nitride to .5 and .5 in the middle of this layer and to 0. and 0. in the silicon, the transition occuring smoothly [7,48-50]. Alternately, the nitride to silicon transition region can be considered to be constituted of a microscopic mixture of the individual phases of Si, SiO_2 and Si_3N_4, sandwiched between undulating interfaces, where the

SiO_2 and Si_3N_4 concentrations decrease towards the silicon, and the Si and SiO_2 concentrations decrease towards the nitride [7].

Thick oxides (> 80 Å) have amorphous structures, while 20-40 Å oxides have structures different to any of the known crystalline forms of SiO_2, Si_3N_4 and SiC [41,42]. Amorphous oxides consist of a random network of 4, 6 and 7-9 member rings of SiO_4 tetrahedra, joined together by bridging O atoms, with corresponding Si-O-Si bond angles of 120^o, 144^o and $160\text{-}180^o$ [38,50,56]. The normal structure for crystalline SiO_2 consists of 6 member rings with a 144^o bond angle. The additional two structures in amorphous silicon dioxide give rise to strained bonds. The density of 120^o bonds increase towards the Si/SiO_2 interface. The greater the density of strained bonds near this interface, the more sensitive the device is to radiation.

The oxide contains a large number of Si-H groups [57]. The Si concentration decreases in the oxide from both interfaces to a minimum near the midpoint, and the N concentration decreases monotonically in the oxide from the oxide/nitride interface towards the Si/SiO_2 interface [54]. Some Si-Si bonds, Si-H bonds and Si dangling bonds are present at the oxide/nitride interface, and some N is present at the Si/SiO_2 interface in devices where the nitride has been deposited at or above 900^oC due to process dependent penetration of the oxide by the N atoms [44,49,50]. The $\equiv$Si center density in the oxide is a low 10^{14} cm^{-3} [28].

During oxidation, after three Si bonds are satisfied, the fourth takes a little longer to react with O [54]. Therefore, as oxidation proceeds, some Si atoms will be left with unfilled bonds, giving excess Si (Si dangling bonds). In thermal oxides, the Si/SiO_2 interface has the form SiO_x, where x goes from 0 to 2 over a distance of up to 25 Å or about six monomolecular layers from the silicon surface (i.e. a silicon rich region) [9,41,58-61] . The transition region is thicker for steam grown oxides than for thermal oxides, but the Si:O ratio is smaller due to the presence of OH and H_2O [58]. The transition region is smaller for low temperature annealed and LPCVD oxides than for other oxides [30]. The SiO_x region is followed by a strained SiO_2 region, 10-40 Å thick, and then by stoichiometric strain-free SiO_2 [9]. The intermediate oxidation states, Si^{3+}, Si^{2+} and Si^{+}, corresponding to Si_2O_3, SiO and Si_2O respectively, are all present in the transition region [38,39,61,62]. The concentration of SiO is greater than that of Si_2O for <100> silicon, and vice-versa for <111> silicon, while <110> silicon yields an even mixture of the two. All three contain Si_2O_3. Some Si_2O_3 is distributed in the oxide bulk also.

The oxide thickness is uniform to within 5 Å, but the oxide layer has undulations of 3-8 Å with a 30-60 Å periodicity and of less than 3 Å with a 200-500 Å periodicity [41]. The roughness of the oxide layer originates from the roughness of the unoxidized silicon surface. Taubenblatt and Helms [63] report that the interfacial width is independent of the oxide thickness. Johannessen et al. [45] report that thinner oxides have smoother interfaces, while Wager and Wilmsen [64] report that the interface width increases with the oxide thickness up to a saturation value of 25 Å at an oxide thickness of 50 Å.

A comprehensive model for the oxide and the Si/SiO_2 interface has been proposed by Johannessen et al. [45,48]. They state that the SiO_x

in the transition region arises from the uneven penetration of O into the silicon surface region, causing necking and finally breaking of $Si\text{-}Si_4$ groups, which then become embedded in the oxide. These groups, which are called inclusions, have different diameters and are present to a depth of 5-10 Å in the oxide. The uneven penetration also causes parts of the silicon surface to protrude into the oxide to a depth of about 20 Å, giving a total interface width of 34-38 Å. The connective regions between the $Si\text{-}Si_4$ formations (in the substrate and the inclusions) and the $Si\text{-}O_4$ formations (in the oxide) have the form $Si\text{-}(Si_{x-1}O_{5-x})$, x=2,3,4, and contain intrinsic bonding defects such as Si dangling bonds and non-bridging O atoms. Additionally, extrinsic bonding defects due to impurities such as H, Na and Cl occur in the connective regions. The density of the intrinsic defects depends on the oxygen arrival rate.

5.2 Nitride Traps, Oxide/Nitride Interface Traps and Their Origins

Energy levels of nitride and oxide/nitride interface traps have been determined using the TSC, V_{TH} (or V_{FB}) Measurement and Charge Centroid Measurement techniques [2,65-70]. Discrete electron traps around 1.0, 1.5 and 2.2 eV below the nitride conduction band edge for the nitride and around 0.8 and 1.3 eV below the nitride conduction band edge for the oxide/nitride interface have been reported. Discrete hole traps occur around 0.8 eV above the nitride valence band edge. Using the Photocurrent Spectroscopy Technique, Kapoor and Bibyk [71] have measured discrete electron traps between 2.5 and 3.8 eV below the nitride conduction band edge at 0.3 eV intervals. Hole traps are much shallower than electron traps [3]. The density and capture cross-section of the nitride and oxide/nitride interface traps have been measured using the Internal Photoelectric Effect, TSC and Charge Centroid Measurement techniques [2,26,29,51]. A nitride trap density of $0.9\text{-}8 \times 10^{18}$ cm^{-3} with a capture cross-section of $0.1\text{-}2 \times 10^{-12}$ cm^2 and an oxide/nitride trap density of $0.5\text{-}3 \times 10^{13}$ cm^{-2} with a capture cross-section of 5×10^{-14} cm^2 have been reported.

Energy levels of states in the nitride have been determined from energy level calculations using the tight-binding and recursion methods [27,28] and from photopopulation measurements [43]. The $\equiv Si$ centers outnumber the $=N$ centers. Therefore, the $\equiv Si$ centers are either neutral ($\equiv Si^0$) or positively charged dangling bonds ($\equiv Si^+$), while the $=N$ centers are doubly occupied and negatively charged. The Si dangling bonds have states 0.8 eV below the nitride conduction band edge and serve as electron traps, while the $=N^-$ centers, which have states just above the nitride valence band edge, serve as hole traps. The Si-Si and Si-H bonding states occur respectively around 0.1 and 0.5 eV above the nitride valence band edge. Being bonding states, they are less capable of trapping than dangling bonds and provide hopping centers for holes, thereby generating a valence band tail. The states, $=N^0$, N-H bonding, N-H antibonding, Si-H antibonding and Si-Si antibonding, are outside the nitride energy bandgap, and the states, $\equiv Si^0$ and $\equiv Si^-$, are deep in it. Therefore, these states do not serve as memory traps. From measurements with the SAES and Photoelectric Effect techniques, Bailey and Kapoor [51] have determined that electron trapping is strongly dependent on the at.% of Si (i.e. excess Si).

Kirk [46] has proposed a Valence Alternate Pair (VAP) model to account for both electron and hole traps in the nitride. VAP states occur when two trivalent N atoms (normal form in stoichiometric Si_3N_4) give rise to a tetravalent N atom and a bivalent N atom, viz:

$$2N_3 \longrightarrow N_4^+ + N_2^-$$

The N_4^+ and the N_2^- have states close to the nitride conduction band edge and valence band edge respectively. These give rise to two other charge states, N_4^0 and N_2^0 (singly occupied intermediate memory states) and N_4^- and N_2^+ (doubly occupied deep permanent states). The problem with this model is that, barring other states, the electron and hole trap densities must be the same, which they are not.

5.3 Si/SiO_2 Interface States, Silicon Bulk Traps and Their Origins

The Si/SiO_2 interface-state density distribution across the silicon energy bandgap for MOS and MNOS devices has been measured using the CCDLTS, Conductance, DLTS, QSCV and TSC techniques. Generally, a U shaped characteristic in the range of 0.1 to 1.05 eV from the valence band edge with a minimum value between 10^{10} and 10^{13} $eV^{-1}cm^{-2}$ has been reported [2,5,6,9,10,12,30,33,36,56,72-83]. The minimum occurs in the range of 0.45 to 0.85 eV from the valence band edge. In some of these results [10,33,36,79-83], one or more humps in the interface-state density characteristic could be observed. A structure that peaks in the upper bandgap region and trails towards the band edges has been obtained from thermal techniques [18,84-88]. A double-peak structure has been reported by Uranwala et al. [89] with the peaks at 0.4 and

0.7 eV from the valence band edge and by Gray and Brown [90] with the peaks at 0.1 and 1.0 eV. The interface-state density values obtained from the thermal techniques range from 2×10^8 to 6×10^{13} $eV^{-1}cm^{-2}$. Using the DLTS technique, Wang [91] has measured a peak of 10^{12} $eV^{-1}cm^{-2}$ near the valence band edge and humps elsewhere in the bandgap. Johnson et al. [79] report of discrete interface trap levels between 0.25 and 0.4 eV from the valence band edge with a maximum of 3×10^{11} $eV^{-1}cm^{-2}$ at 0.35 eV in unannealed thermal oxide MOS devices. These could be removed by a low temperature H_2 anneal.

The energy of an interface trap is sensitive to chemical inhomogeneities (ions; stretched, bent and broken bonds) around it [6]. Due to the large number of such inhomogeneities, there is a statistical deviation in the energy level of a trap. The effect of all these deviations is that the traps appear to be in an energy continuum.

The sharp increase in the interface-state density towards each band edge results in band tailing. These band tails can be considered as states caused by scattering from defects with energies removed from the band edges [83]. Nicollian and Brews [6] explain the band tails as follows: Large clusters of fixed oxide charges may occur with a low probability in small areas and may pull a localized state from the silicon conduction or valence band into the bandgap. These states are most likely to occur near the band edges, resulting in band tails.

The interface-state density is highest in the <111> direction and lowest in the <100> direction, with a reported ratio of the former to the latter between 5:1 and 10:1 [6,9,90,91]. This is attributed to a higher availability of incomplete Si bonds (Si dangling bonds) in the

<111> direction. The interface-state density is higher for dry unannealed oxides than for dry annealed oxides and is lowest for wet oxides [6]. Deposited (LPCVD) oxides give lower interface-state densities than grown oxides [30]. The interface-state density, although assumed to be independant of temperature, drifts slightly with it in practice [12]. This has been attributed to the dependence of the energy bands on temperature [6]. The interface-state density depends on the amount of water in the oxygen used for the oxide growth [90]. It is reduced by the OH' ions, which passivate the defects at the interface. The measured interface-state density depends on the fixed oxide charge due to the common chemical origin of the states and the charge [80]. Thinner oxides result in higher measured interface-state densities due to the contribution from the oxide/nitride interface traps to the measured capacitance [73].

The depth from the silicon surface to which interface states are present in the oxide has been measured using photothermal probing, ellipsometry and comparisions of theoretical models with electrical measurement data. Nicollian and Brews [6] report this depth to be 5-10 Å, which is of the order of an atomic spacing, while Pierret [92] reports it to be 20 Å. Hamasaki [55] claims that interface states are present only in a monomolecular layer of thickness 1.4-2 Å.

An interface trap is a donor if it is neutral when full and positive when empty and an acceptor if it is neutral when empty and negative when full. The measured interface-state density is the sum of the donor and acceptor interface-state densities. Using comparisions of V_{FB} measurements with theoretical models, Huang [81] and Gray and

Brown [90] have found that the interface states in the upper half of the silicon bandgap are acceptors, while those in the lower half are donors. In this case, none of the interface states is charged when E_F intersects E_{is}. From electrical measurements, Kamada and Yagi [93] also report that the interface states in the upper half of the bandgap are acceptors. Nicollian and Brews [6] report that the interface states in the upper half of the bandgap for thermal oxide devices have been determined from conductance measurements to be donors and that there are no reliable data to determine the type of the states in the lower half of the bandgap. Using G-V measurements with temperature as parameter, Singh and Srivastava [33] report that both donor and acceptor states are found in the upper half of the bandgap.

The capture cross-section of the interface states in the upper half of the silicon bandgap has been measured using the Conductance and DLTS techniques [5,12,77,94]. It has been found to decrease towards the conduction band edge from 10^{-13} to 10^{-19} cm^2 first logarithmically and then exponentially. The rapid decrease in the capture cross-section towards the conduction band edge has been attributed partly to the screening of interface states by majority carriers during accumulation biases, which effectively reduces the capture cross-section of donor type traps in n-silicon devices [6].

There are two types of interface states. They are intrinsic states associated with Si dangling bonds at the silicon surface (defects) and with bond angle distortions (disorders) and extrinsic states associated with impurities [33,78]. The bond angle distortions and the impurity states have a Boltzmann distribution and a Gaussian distribution respectively. The midgap states are intrinsic, while the band edge states comprise both types [33,55,61,79]. From comparisions of theory with admittance measurement data, Ancona [10] has shown that the measured interface-state density is composed of two components - the "bond strain" component and the "dielectric" component. The latter has a U shape, is the dominant component and is greater than the measured interface-state density near the valence band edge, while the former is apparently negative towards this band edge. Flietner and Sinh [78] have characterized the interface states using deconvolution of interface-state density curves of dry and wet oxide devices. From a consideration of activation energies, they have shown that the midgap states are due to Si dangling bonds, and that the band edge states are due to bond angle distortions. Using tight-binding and recursion methods to calculate the energy levels of states, Joannopoulos [95] and Laughlin et al. [96] have shown that the conduction band edge states are due to Si-O-Si bond angle distortions, the midgap states are due to Si dangling bonds at the silicon surface, and the valence band edge states are due to Si-Si bond angle distortions. Si dangling bonds and Si-Si bonds in the oxide give rise to states near both band edges. Johannessen et al. [45,48] have attributed the interface states to the bonding defects in the connective regions at the rough Si-SiO_2 boundary (see Section 5.1.3). They claim that donors are associated with Si dangling bonds and acceptors with non-bridging O atoms.

A high interface-state density has several adverse effects. It gives a higher decay rate of memory charge due to a higher back-tunneling current and reduces the oxide breakdown field, the latter resulting in early degradation of the device. It also gives rise to a significant noise component under an applied a.c. field due to fluctuating trap occupancy. Hence, the interface-state density has to be controlled using appropriate processing steps.

The silicon bulk traps are due to impurities such as Au, Cu, Fe, Mn, Ni and Zn which have states near the silicon midgap. They serve as efficient G-R centers for electrons and holes. Chan et al. [32] have calculated a bulk trap density of 4×10^{13} cm^{-3} for MNOS using a plot of G_{gr} versus $1/\omega$.

5.4 Charges

The reported values for the distance of the nitride charge centroid from the oxide/nitride interface range from one third to half the length of the nitride [2,25,29]. The charge distribution extends almost the entire length of the nitride. The location of the charge centroid depends on the magnitude of the charging pulse and is deeper in the nitride for holes than for electrons. The initial nitride charge distribution for a uniform nitride trap density of N_t can be described by [65]

$$n(x) = N_o\exp(-x/\lambda) \qquad (5.1)$$

where x is the distance from the oxide/nitride interface; N_o ($< N_t$) is the charge density at this interface, which depends on the number of injected charges; and λ is the mean free path for trapping of charges in the nitride during injection, which is inversely proportional to N_t. The charge distribution is illustrated in Figure 5.1. The total

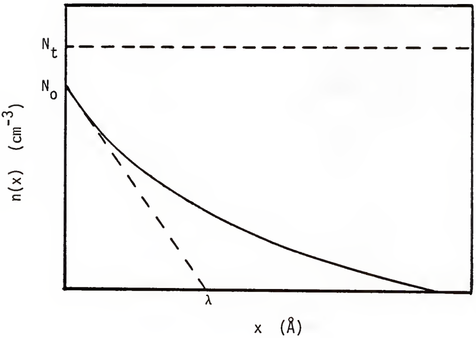


Fig.5.1 Initial nitride charge distribution (theoretical)

trapped charge is $AqN_0\lambda$, where A is the device area [70]. Lehovec [13] reports that the steady state nitride charge distribution depends on the nitride field and β_{PF} but not on the nitride trap depth or density, except in a narrow region close to the oxide/nitride interface.

Oxide charges differ from the Si/SiO_2 interface-state charge in that they do not communicate with the silicon surface. The types of oxide charges are

(i) Fixed oxide charge due to localized charge centers in the oxide near the Si/SiO_2 interface arising from intrinsic defects in the oxide. This charge is a net positive charge, has a density of 10^{10} to 10^{12} cm^{-2} [6,55,61,73] and occurs within 10-30 Å of the silicon surface [9,55]. It has a chemical origin similar to that of the Si/SiO_2 interface states. It is highest in the <111> direction and lowest in the <100> direction [9]. From the point of view of the model of Johannessen et al. described in Section 5.1.3, the fixed oxide charge is associated with the bonding defects in the connective regions between the Si inclusions and the sorrounding oxide. This charge has also been ascribed to the bonding defects in the oxide near the Si/SiO_2 interface whose energy levels lie outside the silicon bandgap but within the oxide bandgap [55]. Raider and Berman [61] attribute this charge to $(Si\text{-}O)^+$ resulting from the termination (by quenching) of the oxidation reaction at the intermediate stage shown below.

$$Si\text{-}Si + (O) \rightarrow \equiv Si^{o} + (Si\text{-}O)^{+} + e \qquad \text{(intermediate)}$$

$$\rightarrow Si\text{-}O\text{-}Si \qquad \text{(complete)}$$

(ii) Mobile oxide charge due to positive ions (mainly Na^+ and other alkali) from the sorrounding system that contaminate the oxide during high temperature processes such as oxidation and annealing. These ions migrate under conditions of high field and/or high temperature and can create a detrimental surface inversion layer in n-channel MNOS FET's and avalanche breakdown in p-channel MNOS FET's due to a reduced depletion width around the p^+-diffusions and an increased field. A mobile oxide charge can cause hysterisis of the C-V curves when the density of this charge exceeds 10^{10} cm^{-2}.

(iii) Trapped oxide charge created by ionizing radiation, hot electron or hole injection from the silicon under an avalanche field and carrier emission by photoemission. This too is a net positive charge.

The three oxide charges can be distinguished by Bias-Temperature (B-T) experiments. Fixed oxide charge does not alter the C-V curves after B-T stress. Mobile oxide charge shifts the C-V curves in the negative voltage direction for positive B-T stress due to the movement of the charge from the nitride side of the oxide towards the Si/SiO_2 interface. It does not affect the C-V curves for negative B-T stress. The trapped oxide charge causes the C-V curves to move in the positive voltage direction after either positive or negative B-T stress due to the charge annealing out.

The fixed oxide charge is reduced by post-oxidation inert gas (e.g. Ar) annealing [55,61]. This may be due to the completion of the oxidation reaction. The mobile ion concentration can be kept low by clean processing [8]. The trapped oxide charge can be reduced by photoemission of electrons from either the metal or the silicon [8], which recombine with the trapped holes.

5.5 Measured Si/SiO$_2$ Interface-State Density Curves and Their Dependence on Processing Variations

Figure 5.2 shows the measured interface-state density curves. Device #15, where the pre-nitride NH$_4$OH dips were not employed, has the highest interface-state density. The NH$_4$OH reduces the interface-state density by etching out regions of agglomoration of impurities and incomplete Si-O bonding (silicon-rich regions), similar to the HCl pre-nitride residual treatment [60], which gives a rougher but cleaner oxide impervious to N. The high ammonia/DCS ratio used for device #4 results in the incorporation of O into the nitride [7,54]. It also results in the penetration of the thin oxide by O and N atoms. Hence, the interface-state density for this device is higher than for devices #3 and #7. As described in Section 5.3, the midgap states are due to Si dangling bonds at the silicon surface, and the conduction band edge states are due to Si-O-Si bond angle distortions. In an H$_2$ anneal, the H passivates the Si/SiO$_2$ and oxide/nitride interface states [4,45,93, 97,98]. Therefore, the higher interface-state density of device #7 compared to device #3 near the conduction band edge can be attributed to an increase in the Si-O-Si bond angle distortions due to the high temperature H$_2$ anneal, and its lower interface-state density near the midgap can be attributed to the passivation of the Si dangling bonds by H during the anneal.

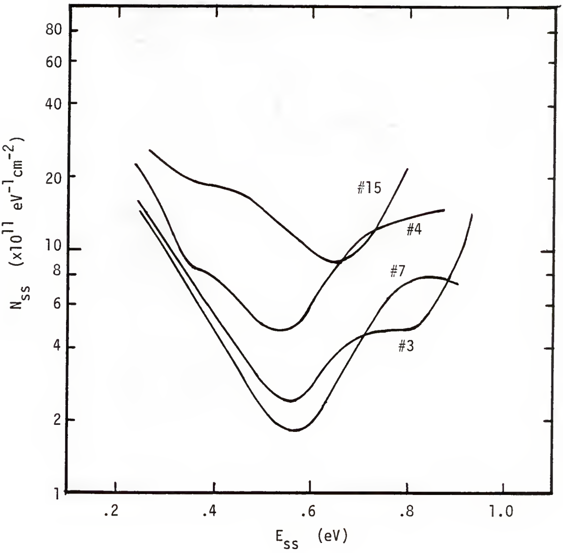


Fig.5.2 The Si/SiO_2 interface-state density versus energy curves for the p^+-gridded MNOS capacitors processed under four different conditions.

CHAPTER VI
RETENTION

6.1 Introduction

The retention time of a nonvolatile MNOS memory device is the time the Write and Erase thresholds remain distinguishable from the non-memory threshold. There are several mechanisms for the decay of stored memory charge at zero bias and under low Read voltages. They are [2,67,99-103]

(i) Direct Tunneling or back-tunneling of electrons from the nitride traps through the oxide to the silicon conduction band.

(ii) Hole injection from the silicon across the oxide into the nitride valence band. This is negligible, except when the stored charge is large, or when a bias is applied. The latter is significant, as it limits the Read time and the allowable number of Read pulses.

(iii) Redistribution of electrons deeper into the nitride or hole conduction towards the oxide/nitride interface.

(iv) Thermal Excitation (TE) of charges at high temperatures due to Poole-Frenkel detrapping and retrapping. The detrapping field is the field due to the stored charge or the low Read voltage.

In reported retention characteristics, the decay rate is linear with log time [4,72,99,100,104]. This behavior has been attributed to the exponential dependence of the tunneling time constant on the distance of traps from the oxide/nitride interface [105]. According to Dobbs et al. [102], the decay rate is slower than log time for negative stored charge. This has been attributed to the negative stored charge being small, which causes the detrapping field to be small. Hezel [47] reports that decay rate is initially faster than log time. Jones and Brown [103] report that a large percentage of charge is lost within 5 minutes of Write (or Erase), while Ferris-Prabhu [19] reports that 20% of the charge is lost in a few microseconds. The decay rate of positive stored charge is greater than that of negative stored charge [4,72,103,104]. This is attributed to the electric field (due to the stored charge) being greater for the former. The application of an external voltage (e.g. a Read pulse) during charge storage increases or decreases the decay rate depending on the polarity of the pulse [65,106]. This is attributed to the Poole-Frenkel barrier lowering effect and its reverse. The decay rate returns to normal when the external voltage is removed.

Initial decay is said to begin at a time t_d, whose reported values range from 10^{-4} to 4 s [65,67,68,105,107]. The equation for decay is written as [65]

$$V_{TH} = \begin{cases} V_{TI} & t \leqslant t_d \\ V_{TI}[1-r\log(t/t_d)] & t \geqslant t_d \end{cases} \qquad (6.1)$$

where V_{TI} is the initial value of V_{TH}, and r is the constant of decay. The initial decay is primarily due to back-tunneling [68]. A break point in the decay curve has been observed at a time of a few seconds to 10^6 secs after Write (or Erase) [68,98,105,108]. The decay rate is higher beyond the break point. This increased decay has been attributed to the redistribution of charge into the nitride and back-tunneling of charges deeper in the nitride which have been detrapped through Poole-Frenkel emission [98,105]. The detrapped charges accumulate at the oxide/nitride interface, creating an additional field, which enhances back-tunneling. Hence, the field assisted back-tunneling time constant of these charges is much less than that associated with their trap locations in the nitride. Therefore, the decay rate is greater.

At high temperatures (> 85°C), TE causes the decay rate to be faster than log time [1,67,100,102,103]. The break point is advanced with increased temperature due to detrapping of charges from deeper traps [100,105]. Sutton [2] reports that the nitride traps around 1.5 eV below the nitride conduction band edge correspond to the 125 - 200°C zero field charge emission range. Therefore, the MNOS has good retention characteristics up to about 150°C. The traps around 1.0 eV below the nitride conduction band edge empty at room temperature. This accounts for the initial rapid decay of stored charge.

6.2 Effects of Processing Variations on the Retention Characteristics

The control of the H concentration and its bonding configurations by varying the nitride deposition conditions and by post-deposition treatments enables to tailor the retention characteristics of the device as desired. The Si-H (or H) concentration determines the positive stored charge, and the amount of unpassivated Si dangling bonds (or excess Si) determines the negative stored charge [43,97]. Thus either n-channel or p-channel operation can be achieved by controlling the Si-H and Si dangling bond concentrations in the nitride. The measured net stored charge is a result of the cancellation of fields due to the positive and negative stored charges. Thus, to obtain an appreciable change in the stored charge, a change of at least a 1 at.% in the relative concentrations of trapping centers is required [43].

Post-nitridation H_2 anneal enables to change the H content and the Si-H/N-H ratio as desired. If the anneal temperature is equal to or greater than the deposition temperature, H is lost, and the Si-H and N-H concentrations decrease [40,43,44,52,97,109]. Hence, the positive stored charge decreases. The H loss in an H_2 anneal is less than in an N_2, inert gas or vacuum anneal. The H loss is diffusion limited and therefore depends on the nitride thickness. Due to the H loss, Si an N dangling bonds are created, which in turn may give rise to Si-N bonds, Si-Si bonds, weak N-N bonds and VAP's [40]. The evolution of these VAP's can be described by

$$2Si_3^0 \rightarrow Si_3^+ + Si_3^-$$

$$2N_2^0 \rightarrow N_2^+ + N_2^-$$

Also, when the anneal temperature exceeds the deposition temperature, Poole-Frenkel type centers are created just below the nitride conduction band edge [107], and the nitride contains a higher percentage of negative charges possibly due to the decomposition of N-H bonds [47]. The Poole-Frenkel type centers increase the nitride conductivity. Hence, retention is reduced. The decay rate decreases with the H_2

anneal temperature up to the deposition temperature and increases thereafter [98]. The latter is attributed to an increase in the nitride conductivity and to enhanced redistribution of charge in the nitride. The decay rate is greatest for unannealed devices.

A higher ammonia/DCS ratio results in a lower at.% of Si [51], a higher nitride trap density [26] and a higher concentration of O impurity from ammonia gas contamination [7,54]. The lower at.% of Si gives a lower negative stored charge. The additional nitride traps are mostly shallow traps and therefore discharge easily [67]. The higher O concentration gives a higher electron trap density [3]. Low ammonia/DCS ratios give silicon rich nitrides with more N-H and N-N bonds, the N atom being coordinated by fewer than three Si atoms [4,40,67]. These result in a high nitride conductivity and a low retentivity.

6.3 Measured Retention Characteristics and Their Dependence on Processing Variations

Figure 6.1 shows the retention curves for device #15, where the pre-nitride NH_4OH dips have not been used. The curves for the other devices are similar. V'_{TH} is linear with log time for both positive and negative stored charges (denoted respectively by +Q and -Q). Circles indicate break points in the decay rate. Table 6.1 shows the average retention data for the four devices. 1st V/dec and 2nd V/dec are the decay rates per decade of time before and after the break point respectively. The initial decay can be attributed to field (due to the stored charge) enhanced back-tunneling of the stored charge. The slower decay rate at larger times suggests that, after some critical

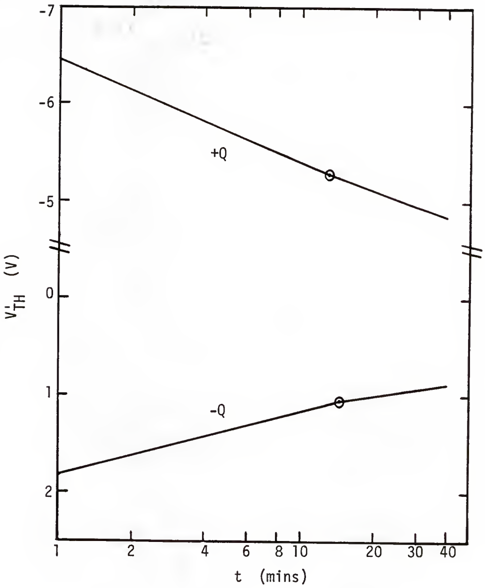


Fig.6.1 Retention curves for the MNOS device #15 for positive stored charge (+Q) and negative stored charge (-Q).

Table 6.1 Average Retention Data for the Four MNOS Devices

Device #	Window (V)		Window Center (V)	1st V/dec		2nd V/dec	
	1 min	40 min	1 min	+Q	-Q	+Q	-Q
3	12.08	9.95	-4.01	.75	.60	.65	.52
4	11.39	9.54	-4.52	.64	.53	.64	.43
7	11.45	9.40	-4.03	.65	.64	.62	.43
15	8.49	5.65	-3.44	1.14	.73	.90	.49

value of the field corresponding to the break point, the decay process is governed by slower processes such as redistribution of the charge into the nitride and large time-constant decay of the charge located deeper in the nitride via back-tunneling. The decay rate is greater for positive stored charge than for negative stored charge. This is due to a greater electric field resulting from the greater positive stored charge and to shallower states being available for hole discharge than for electron discharge.

Device #3 has the largest window. This is because the low ammonia/DCS ratio gives a silicon-rich nitride [4,67], resulting in a large number of Si dangling bonds, which serve as electron traps, and Si-Si centers, which serve as hole traps [28]. However, the availability of shallow traps result in slightly higher decay rates compared to devices #4 and #7. Device #4 has the most negative memory window. This may be due to the formation or inclusion of some positive charge species such as Si-H when a high ammonia/DCS ratio is used. The H_2 anneal passivates Si dangling bonds and $=N^-$ centers (hole traps [28]) and breaks Si-Si bonds to form Si-H and N-H bonds. Thus, the memory window is reduced at both ends for device #7 compared to device #3. For device #15, where the oxidized silicon surface has not been treated, the highest decay rates and the smallest window are obtained. This is due to the high Si/SiO_2 interface-state density of this device and its poor quality oxide, which degrades its retention capability. The deviations in the amount and the decay rate of stored charge of device #15 from those of the other devices are more significant for positive stored charge than for negative stored charge. This suggests that the untreated oxide provides hopping centers for hole transport through the oxide.

Figures 6.2 - 6.5 illustrate the dependence of the retention characteristics on temperature. The window widths decrease with temperature. This is due to the TE process. The decrease is least for device #7 and is only slightly more for device #4. Device #15 has the highest decrease. The window centers become more positive with temperature. This is due to the shallower hole traps being more easily discharged by TE than the deeper electron traps. The movement of the window center is the least for device #7 and the most for device #15, while it is approximately the same for devices #3 and #4. The decay rates for both positive and negative stored charges increase with temperature. This too is due to the TE process. The increase in the decay rate is more significant for positive stored charge than for negative stored charge. This is due to the hole traps being shallower and more numerous than the electron traps. The increase in the positive stored charge decay rate for device #3 is greatest around room temperature, while those for devices #4 and #7 are greatest in the 75 - 125°C range. These suggest that the hole traps in device #3 are shallow, while those in devices #4 and #7 are at an intermediate depth. The hole trap depth of device #15 cannot be correlated to the decay rate due to the influences of a high Si/SiO_2 interface-state density and a poor quality oxide, which tend to hide the detrapping effects of temperature. Negative stored charge decay rate increases significantly after 125°C for devices #3 and #15, indicating deep electron traps in these devices.

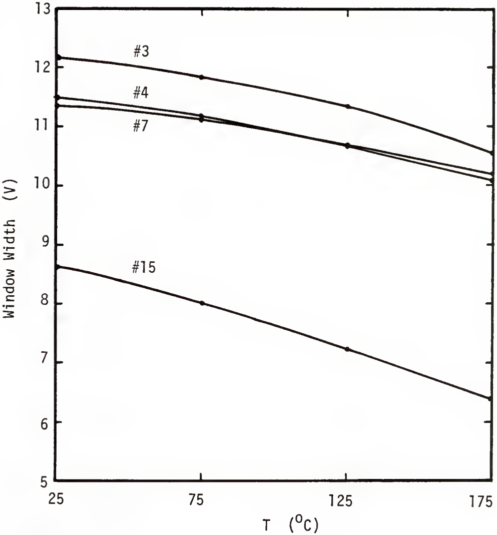


Fig.6.2 Variation of memory window width with temperature for the four MNOS devices.

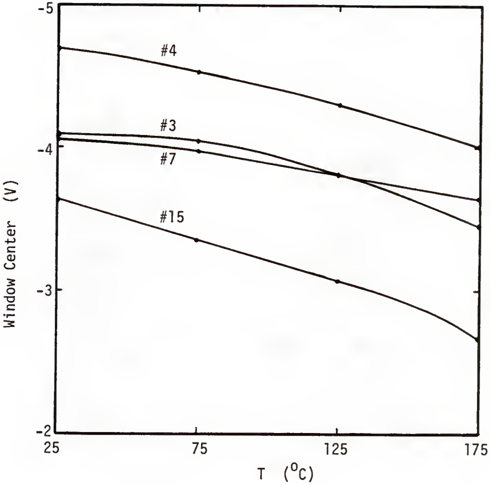


Fig.6.3 Variation of memory window center with temperature for the four MNOS devices.

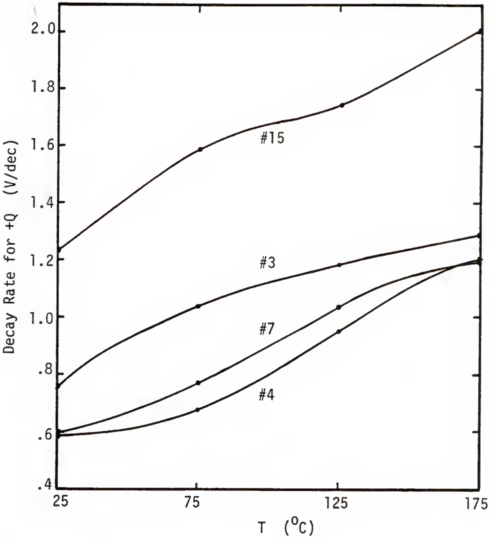


Fig.6.4 Variation of positive stored charge decay rate with temperature for the four MNOS devices.

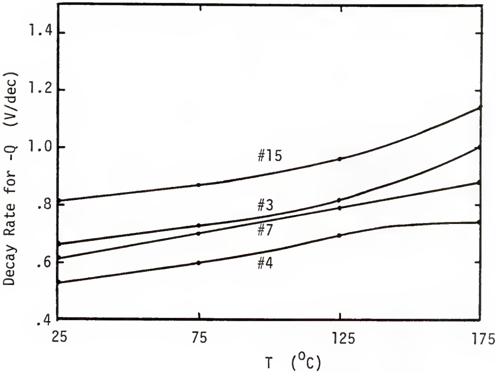


Fig.6.5 Variation of negative stored charge decay rate with temperature for the four MNOS devices.

The good temperature behavior of devices #4 and #7 can be attributed to nitride memory traps created by the high ammonia/DCS ratio in the former and the high temperature H_2 anneal in the latter. The high ammonia/DCS ratio gives some O impurity atoms (electron traps) [3] and possibly some Si-H bonds (hole traps) in the nitride, while the high temperature H_2 anneal gives Si dangling bonds (electron traps) and Si-Si bonds (hole traps) in the nitride [40]. The more significant dependence of the retention characteristics of device #15 on temperature compared to the other devices is attributable to its high Si/SiO_2 interface-state density and its poor quality oxide resulting from the absence of the pre-nitride NH_4OH clean process.

CHAPTER VII
ENDURANCE

7.1 Introduction

The endurance of an MNOS memory device is defined in terms of the number of Write/Erase cycles it can withstand before undergoing complete degradation. Degradation effects in MNOS devices have been studied using Write/Erase cycling and B-T stress. The degradation effects on memory properties are

(i) A reduction in the Write/Erase window width [108].

(ii) An increase in the decay rates [37,105,106,108].

(iii) An overall movement of the threshold memory window in the negative voltage direction [106,108].

Other degradation effects are an increased Si/SiO_2 interface-state density [82,106,108], an increased nitride conductivity [108,109], a broader and shallower quasi-static C-V curve [57,84], a movement of the nitride charge centroid deeper into the nitride [103,105,108], an advancement of the break point in the decay rate [98,105,108] and the breakdown of the thin oxide [3]. It has also been reported that the window width and the decay rates increase with Write/Erase cycling initially and then decrease [104,108].

The decrease in the window width, the increase in the decay rates and the advancement of the decay rate break point are attributed to the increase in the nitride conductivity. The increase in the decay rates is also due to the increase in the Si/SiO_2 interface-state density, which enhances back-tunneling of charge through the oxide. A greater increase in the decay rates has been observed in devices where hydrogen has been incorporated into the oxide (e.g. HCl oxide) than in other devices [18,108]. This suggests that the increase in the Si/SiO_2 interface-state density is due to the breakup of Si-H bonds at the interface. The increase in the nitride conductivity is attributed to the breakup of N-H bonds, which increases the nitride shallow trap density. The negative shift in the memory window is attributed to the creation of donor type Si/SiO_2 interface-states and trapping of holes in the oxide. The breakdown of the oxide is said to be due to forcible hole injection through the oxide. The oxide back-tunneling current density can be written as [108]

$$J_b \propto Q_I N_{ss} e^{-A_1 x_t} \tag{7.1}$$

where Q_I is the initial stored charge, A_1 is a constant, and x_t is the effective tunneling distance from the charge centroid of Q_I. Hence, the initial increase in the decay rates reported by White et al. [108] is attributed to the increase in N_{ss}, and the subsequent decrease in the decay rates is attributed to the increase in x_t resulting from the increase in the nitride conductivity.

Similar to Write/Erase cycling, B-T stress also increases the Si/SiO_2 interface-state density [76,81,89,110] and causes the threshold memory window to move in the negative voltage direction [57,76]. The former is attributed to trapping of electrons in the oxide and to

H transport and accumulation near the Si/SiO_2 interface, while the latter is attributed to an increase in the fixed oxide charge and the stored positive charge. The increase in the stored positive charge is due to an increase in the hole conductivity in the oxide. Discrete Si/SiO_2 interface traps have been reported by Johnson et al. [79] in thermal oxide devices subjected to B-T stress. A greater increase in the Si/SiO_2 interface-state density is obtained for positive B-T stress than for negative B-T stress [111]. This is attributed to a positive species in the oxide being driven to the Si/SiO_2 interface during positive B-T stress. The decay rate of positive stored charge increases under negative B-T stress, while that of negative stored charge remains unchanged [72]. This is attributed to an increase in the donor Si/SiO_2 interface states near the silicon valence band edge.

The performance of an MNOS device is measured in terms of the endurance-retention product, which is the product of the number of Write/Erase cycles and the cumulative retention time. Typical values for this product are 10^{15} cycle-secs for commercial devices and 10^{16} cycle-secs for laboratory devices [1,2,105]. The values are lower when the number of cycles exceed 10^5 [105].

Endurance can be improved by a post-nitridation H_2 anneal, which converts some of the Si dangling bonds and Si-Si bonds in the oxide and at the Si/SiO_2 interface to Si-H bonds, reducing the Si/SiO_2 interface-state density [50,98]. However, the breakdown field of the nitride is reduced by the H_2 anneal [47]. Degradation effects can be removed by a low temperature anneal [76,98,105]. The H lost from the N-H does not diffuse out of the nitride at room temperature [52]. The anneal facilitates the reformation of the N-H.

7.2 Degradation Models

Degradation effects have been explained by Pryor [112] in terms of changes in three types of states that occur in amorphous silicon dioxide and amorphous silicon nitride. These states are

(i) Localized states (defects, dangling bonds, etc.) which can be occupied by carriers only in the vicinity of a single atom with no communication between adjacent states.

(ii) Delocalized states, which comprise a number of localized states between whom communication has been made possible.

(iii) Extended states, which enable carriers to move throughout the volume of the solid.

When the MNOS is subjected to Write/Erase cycling, the high field stresses result in defect related local structural irreversibilities, which cause shifts in local energy levels. Hence, localized states are converted to delocalized states. Further Write/Erase cycling results in a network of delocalized states. These ultimately give rise to extended states, resulting in communication between the substrate and the gate.

Kirk [46] explains the degradation effects in the nitride due to Write/Erase cycling in terms of VAP states. Some singly occupied (intermediate) states become doubly occupied (deep) states. Since the latter are permanently charged, the device becomes degraded. The states near the oxide/nitride interface degrade sooner than those further away. Hence, the charge centroid moves deeper into the nitride due to the decrease in the number of intermediate (memory) states near the oxide/nitride interface. The overall reduction in the number of memory states result in a lower stored charge. The memory properties can be restored by annealing in air, which removes the oxide trapped charges and converts the deep traps back to intermediate traps.

Wurzbach and Grunthaner [50] postulate that Write/Erase cycling causes the breakup of Si-Si and Si-H bonds near the oxide/nitride interface [50], giving rise to Si dangling bonds. Further Write/Erase cycling results in the Si dangling bonds losing their electrons due to impact ionization by injected electrons. Hence, a net positive charge is created, which gives a negative memory window shift. A Si dangling bond is amphoteric. That is, it can donate an electron to create a positive charge, and its singly occupied dangling orbital can also accept an electron. In the latter capacity, it provides a hopping site for electrons to tunnel out of the silicon nitride bulk. Hence, the creation of Si dangling bonds increases the nitride conductivity and thereby decreases its retentivity. Although Si-N and Si-O bonds in silicon nitride and silicon dioxide are stronger than the Si-Si and Si-H bonds, they too break at times, providing amphoteric sites for charge. The Si_3^+ and non-bridging O amphoteric sites from the breakup of Si-O bonds are self annealing, while the amphoteric traps from the breakup of Si-N bonds can be annealed out.

Lai [56] has proposed a model to explain the traps created in the oxide and at the Si/SiO_2 interface by electrical stresses. The breakup of one Si-O bond by a hole is said to relax another adjacent strained Si-O bond. The electrostatic energy is increased, even though the structural energy is reduced. The coulombic charge center, thus created, easily captures an electron to yield a weak bond between the Si and the O or a Si dangling bond and a non-bridging O (dangling bond). This gives a neutral electron trap center in the oxide or at the Si/SiO_2 interface in the upper half of the silicon bandgap.

Jeppson and Svensson [76] describe a model for degradation under negative B-T stress, in which some of the many weakly bonded Si-H bonds in the oxide and at the Si/SiO_2 interface break due to the stress. The H atom, thus released, reacts with the silicon dioxide to form an OH bonded to a Si atom in the oxide. The Si-OH diffuses rapidly in the oxide under negative B-T stress, leaving a Si dangling bond. The increase in the Si/SiO_2 interface-state density is attributed to the creation of these dangling bonds at the interface. In the case of high fields, oxide hole traps become aligned with the silicon valence band, resulting in hole trapping in the oxide. These trapped holes also give rise to traps at the Si/SiO_2 interface. Annealing reverses these processes and hence reduces the Si/SiO_2 interface-state density.

Degradation mechanisms under positive B-T stress have been described by Suzuki et al. [57] and Huang [81]. Electron injection from the silicon breaks strained Si-O bonds and Si-H bonds near the Si/SiO_2 interface, giving rise to Si dangling bonds. Also, holes injected from the gate traverse the length of the nitride due to the high density of states near the nitride valence band edge and break some of the Si-H present in the oxide and at the Si/SiO_2 interface, giving rise to Si dangling bonds and trivalent silicon. The Si dangling bonds at the

Si/SiO_2 interface increase the interface-state density, while the trivalent Si atoms in the oxide act as hole traps. A microchannel of Si-H bonds along the oxide can lead to a conductive path through the oxide after extended stress. Therefore, devices with high hole conductivity in the nitride degrade easily.

7.3 Experimental Results of Degradation Studies

Figure 7.1 shows the variation of the interface-state density versus energy curve with Write/Erase cycling for a p^+-gridded MNOS device. It can be seen that this curve becomes higher and broader with the number of Write/Erase cycles. The increase in the interface-state density can be attributed to the creation of Si dangling bonds, while the broadening of the curve up to $4x10^9$ cycles can be attributed to the relieving of the strain in the Si-O-Si and Si-Si bonds at the silicon surface by the electrical stresses and to the annihilation of the band edge states (Si dangling bonds and Si-Si bonds in the oxide) by the repeated change in polarity of the electric field. Figure 7.2 shows the surface potential versus gate voltage curves corresponding to the interface-state density curves of Figure 7.1. It can be seen that the curve becomes steeper up to $4x10^9$ cycles and then becomes shallower. The former is due to the reduced band edge states, which results in a lower screening charge and therefore a greater amount of band-bending, and the latter is due to the high interface-state density at 10^{10} cycles, which gives rise to a large screening charge and therefore reduced band-bending.

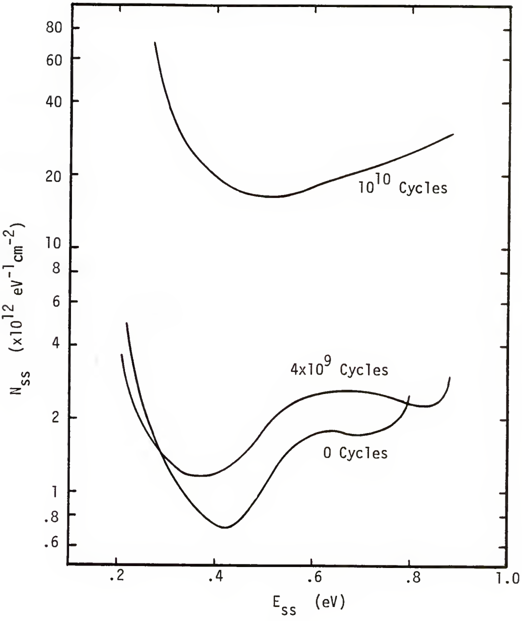


Fig.7.1 Variation of the Si/SiO_2 interface-state density versus energy curve with Write/Erase cycling for the p^+-gridded MNOS capacitor.

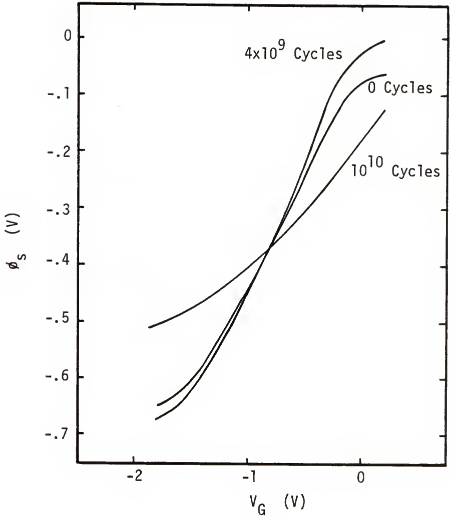


Fig.7.2 Variation of the surface potential versus gate voltage curve with Write/Erase cycling for the p^+-gridded MNOS capacitor.

Figures 7.3 - 7.6 show the variation of the interface-state density versus energy curves with Write/Erase cycling for the four MNOS devices, and Figures 7.7 - 7.10 show the variation of these curves with negative B-T stress time. A general pattern of the curves becoming higher and broader due to the stresses can be observed. The bandgap humps tend to be ironed out and pulled towards the band edges. This is due to the saturation of the impurity sites and the relieving of the strained bonds. The effect of stress on the four devices seem to be similar, except that device #15 undergoes significantly less change than the others. This is due to the poor quality nature of the oxide in device #15, which allows carriers to tunnel through it, reducing the damage at the Si/SiO_2 interface due to carrier pile-up. Comparing the two sets of curves, it can be seen that Write/Erase cycling and negative B-T stress produce similar effects at the Si/SiO_2 interface.

Figures 7.11 - 7.14 show the variation of the retention characteristics with number of Write/Erase cycles and hours of negative B-T stress for the four MNOS devices. The window widths increase with number of Write/Erase cycles but decrease with B-T stress time, except for device #15, where the window width increases with stress time initially and then decreases. The increase in the window widths with Write/Erase cycling can be attributed to the creation of hole traps in the nitride. The window width has been shown [37] to decrease with extended Write/Erase cycling, and therefore it can be expected to decrease after 10^7 cycles. The decrease in the window widths with B-T stress time is due to the hole traps in the nitride being annihilated

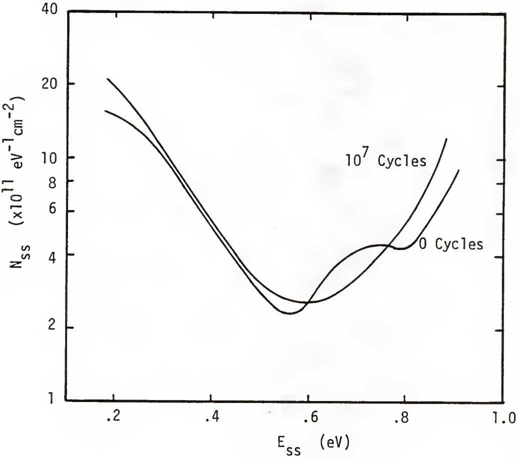


Fig.7.3 Variation of the Si/SiO_2 interface-state density versus energy curve with Write/Erase cycling for device #3.

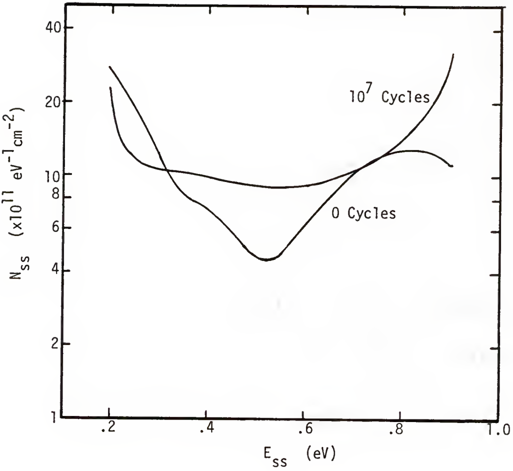


Fig.7.4 Variation of the Si/SiO_2 interface-state density versus energy curve with Write/Erase cycling for device #4.

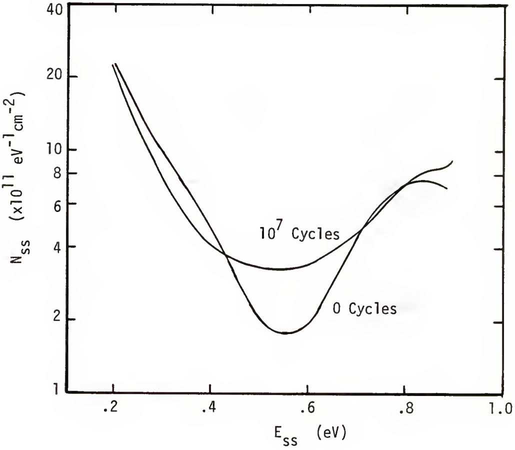


Fig.7.5 Variation of the Si/SiO_2 interface-state density versus energy curve with Write/Erase cycling for device #7.

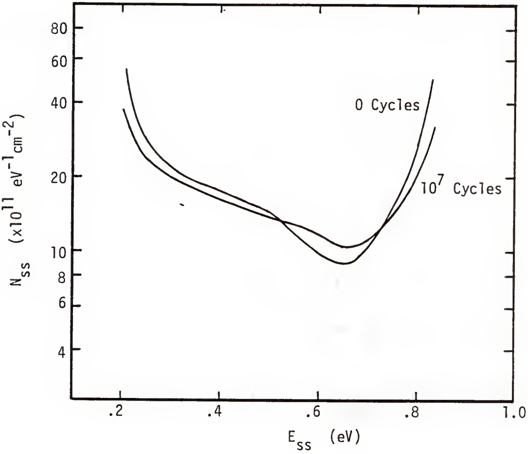


Fig.7.6 Variation of the Si/SiO_2 interface-state density versus energy curve with Write/Erase cycling for device #15.

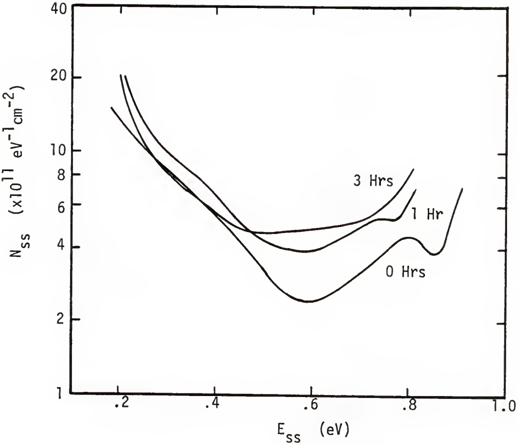


Fig.7.7 Variation of the Si/SiO_2 interface-state density versus energy curve with negative B-T stress time for device #3.

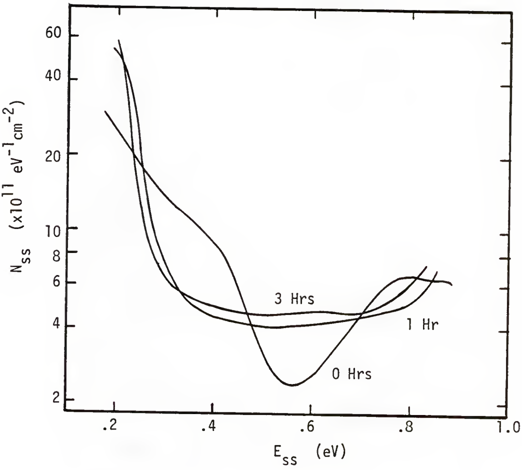


Fig.7.8 Variation of the Si/SiO_2 interface-state density versus energy curve with negative B-T stress time for device #4.

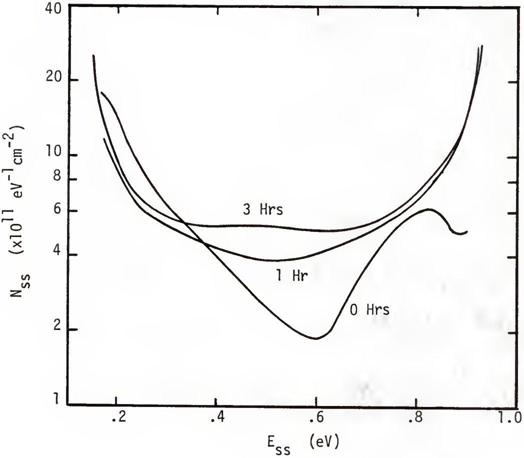


Fig.7.9 Variation of the Si/SiO_2 interface-state density versus energy curve with negative B-T stress time for device #7.

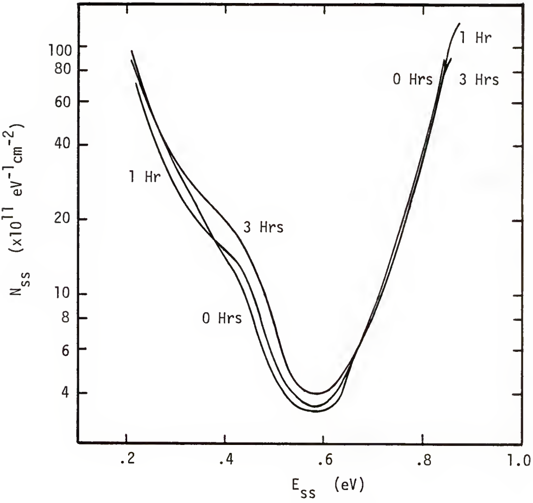


Fig.7.10 Variation of the Si/SiO_2 interface-state density versus energy curve with negative B-T stress time for device #15.

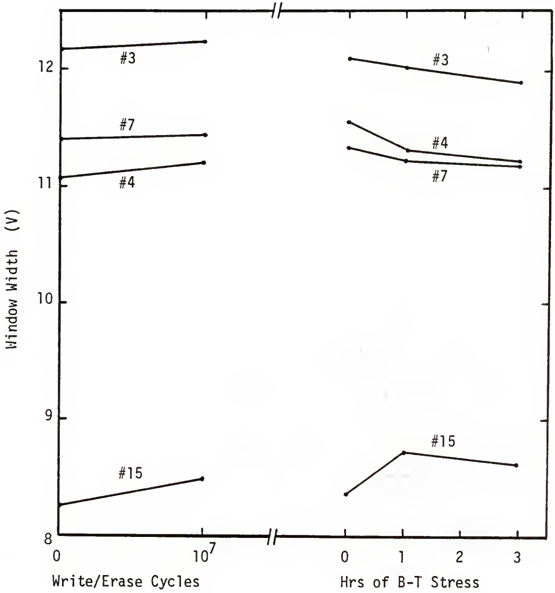


Fig.7.11 Variation of memory window width with number of Write/Erase cycles and hours of negative B-T stress for the four MNOS devices.

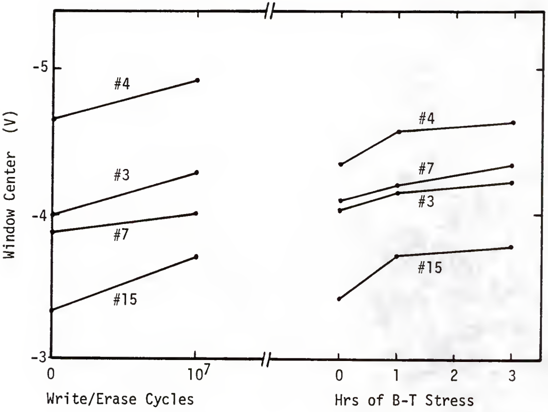


Fig.7.12 Variation of memory window center with number of Write/Erase cycles and hours of negative B-T stress for the four MNOS devices.

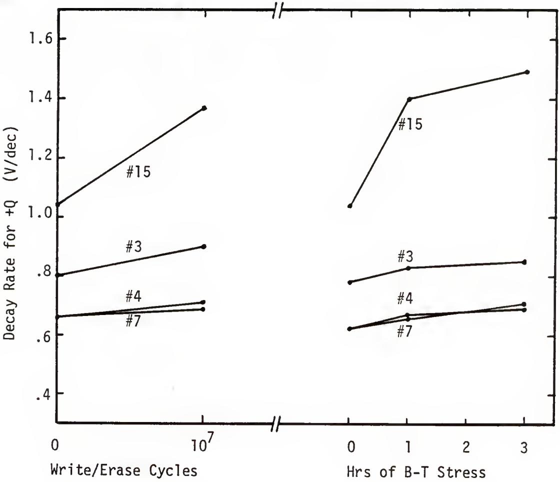


Fig.7.13 Variation of positive stored charge decay rate with number of Write/Erase cycles and hours of negative B-T stress for the four MNOS devices.

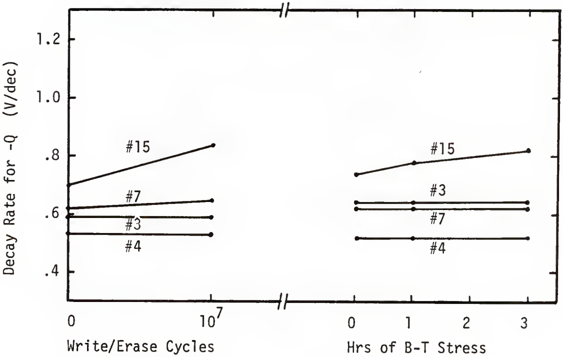


Fig.7.14 Variation of negative stored charge decay rate with number of Write/Erase cycles and hours of negative B-T stress for the four MNOS devices.

by hole injection from the silicon under negative B-T stress. The errant behavior of device #15 is due to the poor quality oxide, which allows enough holes to be injected into the nitride to offset the effects of the annihilation. The window centers become more negative for both Write/Erase cycling and negative B-T stress, consistent with other reported results. The causes for this are an increase in the donor Si/SiO_2 interface states and hole trapping in the oxide. The positive stored charge decay rates increase with number of Write/Erase cycles and negative B-T stress time, while the negative stored charge decay rates remain almost unchanged, except for device #15. The increase in the positive stored charge decay rates can be attributed to the increase in the Si/SiO_2 interface-state density and the increase in the nitride conductivity. The fact that the increase in the negative stored charge decay rate is significantly smaller than that in the positive stored charge decay rate can be attributed to the smaller stored charge and the lower nitride conductivity for electrons than for holes.

The effects of positive B-T stress on the four devices were studied cursorily. They were found to be more detrimental than those of negative B-T stress in that the insulator capacitance broke down sooner. This suggests that holes injected from the gate contribute significantly to the degradation of the insulator layers of MNOS devices and that a greater damage is done in the positive half of the Write/Erase cycle than in the negative.

The increase in the Si/SiO_2 interface-state density of device #15 is not commensurate with the very significant degradation of its retention characteristics. From this, it can be concluded that degradation proceeds through an increase in the nitride conductivity and a deterioration of the oxide as well as through an increase in the Si/SiO_2 interface-state density. Of the four devices, device #7 manifests the best endurance behavior of the retention characteristics. This stresses the importance of the post-nitridation H_2 anneal. The use of the higher ammonia/DCS ratio in device #4 compared to device #3 does not seem to influence the effects of stress on the retention characteristics of the device. Of the four devices, device #15 undergoes by far the worst degradation of the retention characteristics. This is due to the susceptibility of the poor quality oxide to damage by the stresses.

7.4 Quantitative Illustration of Degradation Effects

A few simple calculations and some heuristic assumptions are used to illustrate the individual contributions to degradation from its two major sources - the increase in the Si/SiO_2 interface-state density and the change in the nitride stored charge. The interface-state density versus energy curves of device #4 in Figure 7.4 are used for this purpose.

To start with, it is assumed that the Si/SiO_2 interface states in the upper half of the silicon bandgap are donors and those in the lower half are acceptors. The interface-state charge is given by

$$Q_{ss} = \int A_1 q N_{ss} dE_{ss}$$

$$= A_1 q \times \text{area under the interface-state density curve} \qquad (7.2)$$

where A_1 is the area of the n-region. It was found that the capacitance value used to determine V'_{TH} (for retention measurements) corresponded to 0.214 eV on the energy scale for both 0 and 10^7 Write/Erase cycles. Therefore, at the instant of the V'_{TH} measurement, all the states in the upper half of the bandgap are charged (positively), while only the states up to 0.214 eV in the lower half of the bandgap are charged (negatively). Hence, as the 10^7 cycle curve is higher than the 0 cycle curve in the upper half of the bandgap and lower near the valence band edge in the lower half of the bandgap, the total excess positive charge in the interface states at 10^7 cycles over that at 0 cycles is the summation of the area between the curves in the upper half of the bandgap and the area between the curves up to 0.214 eV in the lower half of the bandgap. These areas were obtained by extrapolating the lower curve vertically to meet the upper curve at each band edge and counting elemental rectangles. The total number of excess positive charges in the interface states after 10^7 Write/Erase cycles was found to be 1.27×10^{11} cm^{-2}. Then, from (7.2), the difference in the interface-state charge, ΔQ_{SS}, works out to 1.11×10^{-11} C.

Assuming that the effect of the series resistance is negligible at 1 MHz (the frequency of the retention measurements), the equivalent circuit comprises of two parallel paths, one through the n-region to the substrate and the other through the p^+-grid to the substrate. That is, the entire gate voltage appears across each of these paths. Hence, in (2.19), to obtain the contribution to V_{TH} from the silicon charges, we need to consider only the path through the n-region. Then, assuming that the effect of the nitride charge centroid movement is negligible up to 10^7 Write/Erase cycles, we obtain from (2.19)

$$\Delta V'_{TH} = \Delta V_{TH} = -\Delta Q_{ss}/C_{In} - \Delta Q_T/C_N \qquad (7.3)$$

where $\Delta V'_{TH}$, ΔV_{TH}, ΔQ_{ss} and ΔQ_T are the increments in the respective quantities after 10^7 Write/Erase cycles. Writing ΔV_{ss} and ΔV_{sc} as the contributions to $\Delta V'_{TH}$ from the increase in the Si/SiO$_2$ interface-state density and the change in the stored charge respectively, we have

$$\Delta V_{ss} = -\Delta Q_{ss}/C_{In} \qquad (7.4)$$

$$\Delta V_{sc} = -\Delta Q_T/C_N \qquad (7.5)$$

Substituting for ΔQ_{ss} in (7.4),

$$\Delta V_{ss} = -0.15\ V$$

From the window center versus Write/Erase cycles curve for device #4 in Figure 7.12, we obtain

$$\Delta V'_{TH} = -0.27\ V$$

Therefore, from (7.3), (7.4) and (7.5),

$$\Delta V_{sc} = -0.12\ V$$

Hence, the contributions from the two degradation mechanisms are comparable, which supports the inference made is Section 7.3.

From (7.5),

$$\Delta Q_T = -\Delta V_{sc} C_N \qquad (7.6)$$

From (7.6), the change in the nitride stored charge works out to a positive charge of 1.29×10^{-11} C. In Figure 7.11, the window width is seen to increase with Write/Erase cycles, and in Figure 7.12, the window center is seen to become more negative with Write/Erase cycles. These suggest that the positive stored charge has increased, which supports the above result. The increase in the nitride trap density is

$$\Delta N_t = \Delta Q_T/(qV_n) \tag{7.7}$$

where V_n is the nitride volume. Hence, ΔN_t works out to 2.3×10^{16} cm^{-3}. This seems reasonable when compared to the reported nitride trap densities of 0.9 - 8×10^{18} cm^{-3} (see Section 5.2).

CHAPTER VIII
CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

The measured interface-state density values can be only as accurate as the model used to extract them using the QSCV technique. The model used in this study, which was theoretically derived, has been adequately verified from the simulation of the C-V curves. A modification of the model was needed to simulate the high frequency C-V curves in accumulation. This did not affect the C-V curves in the depletion and weak inversion regions.

The QSCV technique is an economic and convenient tool to obtain several device parameters, including the Si/SiO_2 interface-state density. The interface-state density characteristic had a general U shape and a minimum between $2x10^{11}$ and 10^{12} $eV^{-1}cm^{-2}$ for the four devices. These values agree well with other reported results. The humps in the interface-state density curves were explained in terms of the nitride deposition process. The four devices in increasing order of magnitude of the interface-state density curves are #7, #3, #4 and #15. The differences were interpreted in terms of the different processing conditions.

A V_{TH} decay which was linear with log time, a break point in the decay rate, with a smaller decay rate at larger times, and a higher decay rate for positive stored charge than for negative stored charge were observed. These were explained in terms of discharging mechanisms and electric fields due to stored charges. Device #3 had the largest memory window, while device #15 had the smallest. However, the decay rates of device #3 were higher than those of devices #4 and #7. The highest decay rates were observed for device #15. Device #4 had the most negative memory window, while device #15 had the least. These retention characteristics were interpreted in terms of the chemical compositions and structures resulting from the different processing conditions. The retention characteristics of the four devices deteriorated uniformly with temperature, while the devices retained their relative values. The variations of the decay rates with temperature were used to identify relative trap depths.

The effects of Write/Erase cycling and negative B-T stress on the Si/SiO_2 interface-state density and retention characteristics of the four devices were studied. Higher and broader interface-state density curves resulted from the stresses. Except for the negative stored charge decay rates for devices #3, #4 and #7 and the window widths under Write/Erase cycling, the retention characteristics deteriorated uniformly with continued stress. The negative stored charge decay rates for the three devices remained almost unchanged by stress, and the window widths increased with number of Write/Erase cycles (up to 10^7). These effects were explained in terms of several degradation mechanisms. Brief studies of the effects of positive B-T stress on the four devices and of Write/Erase cycling on the surface potential versus gate voltage curve were made. It was shown that degradation proceeds through the deterioration of the insulator layers as well as through the increase in the Si/SiO_2 interface-state density, and that the positive half of the Write/Erase cycle is more detrimental to the device than the negative half. All in all, device #7 had the best endurance behavior, which stresses the importance of the post-nitridation H_2 anneal, while devices #3 and #4 had generally the same degradation characteristics. Device #15 had by far the worst endurance characteristics, stressing the importance of the NH_4OH clean process.

8.2 Recommendations

The origin of the space charge loss conductance may be identified through quantum mechanical means if necessary. Also, quantitative descriptions for this conductance and the G-R conductances may be formulated. The actual majority carrier capture cross-section values of the Si/SiO_2 interface states can be obtained by curve fitting of conductance data. Hence, a more complete and accurate simulation of the C-V curves could be obtained.

A few more processing conditions such as a thicker oxide, an H impregnated oxide, additional ammonia/DCS ratios and inert gas and N_2 anneals could be used to obtain a deeper insight into trap origins and trapping and detrapping mechanisms. Measurement of the retention characteristics at more temperature points may yield a more concrete knowledge of trap levels in the nitride. Degradation studies may be extended further to determine whether the observed endurance behaviors will continue under prolonged stress.

REFERENCES

1. J.E. Brewer, "MNOS Retention Defects," *M.Sc. Thesis*, Univ. of Maryland (May 1984).

2. W.G. Sutton, "Charge Trapping in Interface Doped MNOS Structures," *Ph.D. Dissertation*, Univ. Microfilms International, Ann Arbor, Michigan (1981).

3. E. Suzuki, H. Hiraishi, K. Ishii and Y. Hayashi, "A Low-Voltage Alterable EEPROM with Metal-Oxide-Nitride-Oxide-Semiconductor (MONOS) Structures," *IEEE Trans. Electron. Devices*, **ED-30**, 122-128 (Feb. 1983).

4. S. Zirinsky, "Charge Transfer Properties of MNOS Structures as Influenced by Processing Parameters," *J. Electron. Mat.*, **4,#3**, 591-624 (1975).

5. A. Singh and J.G. Simmons, "Determination of Energy Density Distribution and Capture Cross-Section of Interface States in the Metal-Nitride-Oxide-Semiconductor (MNOS) Structure," *Solid-State Electron.*, **25**, 219-226 (1982).

6. E.H. Nicollian and J.R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology*, John Wiley & Sons, Inc., New York (1982).

7. J.S. Johannessen, C.R. Helms, W.E. Spicer and Y.E. Strausser, "Auger Depth Profiling of MNOS Structures by Ion Sputtering," *IEEE Trans. Electron. Devices*, **ED-24**, 547-551 (May 1977).

8. R.S. Muller and T.I. Kamins, *Device Electronics for Integrated Circuits*, John Wiley & Sons, Inc., New York (1977).

9. S.M. Sze, *Physics of Semiconductor Devices, 2nd Edition*, John Wiley & Sons, Inc., New York (1981).

10. M.G. Ancona, "Electric Field Induced Effects at the Si-SiO_2 Interface: Theory and Experiment," *J. Appl. Phys.*, **54**, 5231-5239 (Sep. 1983).

11. K. Lehovec, "Frequency Dependence of the Impedance of Distributed Surface States in MOS Structures," *Appl. Phys. Lett.*, **8**, 48-50 (15 Jan. 1966).

12. H. Deuling, E. Klausmann and A. Goetzberger, "Interface States in Si-SiO_2 Interfaces," Solid-State Electron., **15**, 559-571 (1972).

13. K. Lehovec, "Charge Distribution in the Nitride of MNOS Memory Devices," J. Electron. Mat., **6,#2**, 77-93 (1977).

14. D. Frohman-Bentchkowsky and M. Lenzlinger, "Charge Transport and Storage in Metal-Nitride-Oxide-Silicon (MNOS) Structures," J. Appl. Phys., **40**, 3307-3320 (July 1969).

15. C. Svensson and I. Lundstrom, "Trap-Assisted Charge Injection in MNOS Structures," J. Appl. Phys., **44**, 4657-4663 (Oct. 1973).

16. K.I. Lundstrom and C.M. Svensson, "Properties of MNOS Structures," IEEE Trans. Electron. Devices, **ED-19**, 826-836 (June 1972).

17. N. Gordon and W.C. Johnson, "Switching Mechanism in Thin-Oxide MNOS Devices," IEEE Trans. Electron. Devices, **ED-20**, 253-256 (Mar. 1973).

18. A.V. Ferris-Prabhu, "Charge Transfer by Direct Tunneling in Thin-Oxide Memory Transistors," IEEE Trans. Electron. Devices, **ED-24**, 524-530 (May 1977).

19. A.V. Ferris-Prabhu, "Theory of MNOS Memory Device Behavior," IBM J. Res. Develop., 125-134 (Mar. 1973).

20. H. Yamamoto, H. Iwasawa and A. Sasaki, "Discharging Process by Multiple Tunnelings in Thin-Oxide MNOS Structures," IEEE Trans. Electron. Devices, **ED-29**, 1255-1261 (Aug. 1982).

21. H.E. Maes and R.J. Van Overstraeten, "Low-Field Transient Behavior of MNOS Devices," J. Appl. Phys., **47**, 664-666 (Feb. 1976).

22. J.J. Chang, "Theory of MNOS Memory Transistor," IEEE Trans. Electron. Devices, **ED-24**, 511-518 (May 1977).

23. M.M.E. Beguwala and T.L. Gunckel, "An Improved Model for the Charging Characteristics of a Dual-Dielectric (MNOS) Nonvolatile Memory Device," IEEE Trans. Electron. Devices, **ED-25**, 1023-1030 (Aug. 1978).

24. M. Conti and A. Vanin, "Oxide and Nitride Conduction Mechanisms in MNOS Structures," Thin Solid Films, **14**, 211-220 (1972).

25. B.H. Yun, "Electron and Hole Transport in CVD Si_3N_4 Films," Appl. Phys. Lett., **27**, 256-258 (15 Aug. 1975).

26. H. Schauer and E. Arnold, "Simple Technique for Charge Centroid Measurement in MNOS Capacitors," J. Appl. Phys., **50**, 6956-6961 (Nov. 1979).

27. J. Robertson, "Defect and Impurity States in Silicon Nitride," J. Appl. Phys., **54**, 4490-4493 (Aug. 1983).

28. J. Robertson and M.J. Powell, "Gap States in Silicon Nitride," Appl. Phys. Lett., **44**, 415-417 (15 Feb. 1984).

29. B.H. Yun, "Measurements of Charge Propagation in Si_3N_4 Films," Appl. Phys. Lett., 25, 340-342 (15 Sep. 1974).

30. G. Pananakakis, G. Kamarinos, M. El-Sayed and V. Le Goascoz, "Electrical Characterization of Al-SiO_2-Si (n-Type) Tunnel Structures. Influence of LPCVD and LPO_2 Oxide Growth Technologies on the Properties of the Si-SiO_2 Interface," Solid-State Electron., **26**, 415-426 (1983).

31. H.A. Mar and J.G. Simmons, "Determination of Bulk Trap Parameters Using Thermal Dielectric Relaxation Techniques," Solid-State Electron., **17**, 1181-1185 (1974).

32. K.W. Chan, J.G. Simmons and H.A. Mar, "Small-Signal Response of Bulk Traps in MNOS Devices," Solid-State Electron., **20**, 249-254 (1977).

33. R.J. Singh and R.S. Srivastava, "Investigations of Interface-State Density in Si-MOS Structures," Solid-State Electron., **25**, 227-232 (1982).

34. H.A. Mar, J.G. Simmons and G.W. Taylor, "Frequency Response of Bulk Traps in the Metal-Oxide-Silicon Structure under Strong-Inversion Conditions," Solid-State Electron., **30**, 241-247 (1977).

35. D.J. Hamilton and W.J. Howard, Basic Integrated Circuit Engineering, McGraw-Hill, Inc., New York (1975).

36. P. Tutto and J. Balazs, "Insulator Polarization Effect in Quasi-Static and High-Frequency C(V) Curves," Solid-State Electron., **25**, 45-48 (1982).

37. HARRIS Semiconductor Programs Division, Final Report, CSDL Non-volatile Memory Development, Phase I, "Transfer of Silicon Nitride Deposition Process from Westinghouse to Harris."

38. F.J. Grunthaner, P.J. Grunthaner, R.P. Vasquez, B.F. Lewis, J. Maserjian and A. Madhukar, "High Resolution X-Ray Photoelectron Spectroscopy as a Probe of Local Atomic Structure: Application to Amorphous SiO_2 and the Si-SiO_2 Interface," Phys. Rev. Lett., **43**, 1683-1686 (26 Nov. 1979).

39. G. Hollinges and F.J. Himpsel, "Probing the Transition Layer at the SiO_2-Si Interface Using Core Level Photoemission," Appl. Phys. Lett., **44**, 93-95 (1 Jan. 1984).

40. T. Makino, "Composition and Structure Control by Source Gas Ratio in LPCVD SiN_x," J. Electrochem. Soc., **130**, 450-455 (Feb. 1983).

41. O.L. Krivanek and J.H. Mazur, "The Structure of Ultrathin Oxide on Silicon," Appl. Phys. Lett., **37**, 392-394 (15 Aug. 1980).

42. B. Agins, S. Rigo, F. Rochet, M. Froment, C. Maillot, H. Roulet and G. Dufour, "Structural Evolution of Very Thin Silicon Oxide Films during Thermal Growth in Dry Oxygen," Appl. Phys. Lett., **44**, 48-50 (1 Jan. 1984).

43. P.S. Peercy and H.J. Stein, "Materials Analysis of Silicon Nitride," Electrochem. Soc. Spring Meeting (8-13 May 1983).

44. H.J. Stein, S.T. Picraux and P.H. Holloway, "Analyses for Stoichiometry and for Hydrogen and Oxygen in Silicon Nitride Films," IEEE Trans. Electron. Devices, **ED-25**, 1008-1014 (Aug. 1978).

45. J.S. Johannessen, W.E. Spicer and Y.E. Strausser, "An Auger Analysis of the SiO_2-Si Interface," J. Appl. Phys., **47**, 3028-3037 (July 1976).

46. C.T. Kirk Jr., "Valence Alternation Pair Model of Charge Storage in MNOS Memory Devices," J. Appl. Phys., **50**, 4190-4195 (June 1979).

47. R. Hezel, "High Temperature Annealing of MNOS Devices and Its Effect on Si-Nitride Stress, Interface Charge Density and Memory Properties," J. Electron. Mat., **8,#4**, 459-484 (1979).

48. J.S. Johannessen, W.E. Spicer and Y.E. Strausser, "Auger Depth Profiling of Interfaces in MOS and MNOS Structures," J. Vac. Sci. Technol., **13**, 849-855 (July/Aug. 1976).

49. J.A. Wurzbach, F.J. Grunthaner and J. Maserjian, "Composition and Electron Stress Effects in Silicon Nitride Thin Films Made by Thermal Growth and Chemical Etching of LPCVD MNOS Structures as Studied by X-Ray Photoelectron Spectroscopy (XPS)," J. Vac. Sci. Technol., 20, 962-965 (Apr. 1982).

50. J.A. Wurzbach and F.J. Grunthaner, "Compositional Depth Profile of a Native Oxide LPCVD MNOS Structure Using X-Ray Photoelectron Spectroscopy and Chemical Etching," J. Electrochem. Soc., 130, 691-699 (Mar. 1983).

51. R.S. Bailey and V.J. Kapoor, "Variation in the Stoichiometry of Thin Silicon Nitride Insulating Films on Silicon and Its Correlation with Memory Traps," J. Vac. Sci. Technol., **20**, 484-487 (Mar. 1982).

52. T.P. Smirnova, V.I. Belyi and L.V. Chramova, "The Mechanism of the N-H Bond Decomposition in Silicon Nitride Films," Thin Solid Films, **74**, 287-293 (1980).

53. H.L. Hwang, C.C. Hwu, J.C. Liue and H.H. Lin, "Analysis of Low-Pressure Chemical Vapor Deposited Silicon Nitride by Rutherford Backscattering Spectrometry," Appl. Phys. Lett., **41**, 844-846 (1 Nov. 1982).

54. H. Nakayama and T. Enomoto, "Effects of Nitride Deposition Conditions on Characteristics of an MNOS Nonvolatile Memory Transistor," Japan. J. Appl. Phys., 18, 1773-1779 (Sep. 1979).

55. M. Hamasaki, "Generation Kinetics of Oxide Charges and Surface States during Oxidation of Silicon," Solid-State Electron., 25, 205-211 (1982).

56. S.K. Lai, "Interface Trap Generation in Silicon Dioxide when Electrons are Captured by Trapped Holes," J. Appl. Phys., **54**, 2540-2545 (May 1983).

57. E. Susuki, Y. Hayashi and H. Yanai, "A Model of Degradation Mechanisms in Metal-Nitride-Oxide-Semiconductor Structures," Appl. Phys. Lett., **35**, 790-792 (15 Nov. 1979).

58. R.A. Clarke, R.L. Tapping, M.A. Hopper and L. Young, "An ESCA Study of the Oxide at the Si-SiO_2 Interface," J. Electrochem. Soc., **122**, 1347-1350 (Oct. 1975).

59. W.L. Harrington, R.E. Honig, A.M. Goodman and R. Williams, "Low-Energy Ion-Scattering Spectrometry (ISS) of the SiO_2/Si Interface," Appl. Phys. Lett., **27**, 644-645 (15 Dec. 1975).

60. R.M. McLouski, M.C. Peckerar and J.J. Schreurs, "The Effects of HCl Residual Treatment on MNOS Memory Transistors," J. Electrochem. Soc., 125, 985-988 (June 1978).

61. S.I. Raider and A. Berman, "On the Nature of Fixed Oxide Charge," J. Electrochem. Soc., 125, 629-633 (Apr. 1978).

62. T. Hattori and T. Suzuki, "Si-SiO_2 Interface Structures on Si (100), (111) and (110) Surfaces," Appl. Phys. Lett., 43, 470-472 (1 Sep. 1983).

63. M.A. Taubenblatt and C.R. Helms, "Ion Knock-on Broadening Effects in Auger Sputter Profiling Studies of Ultra Thin SiO_2 Layers on Si," J. Appl. Phys., 54, 2667-2671 (May 1983).

64. J.F. Wager and C.W. Wilmsen, "Auger Analysis of Ultrathin SiO_2 Layers on Silicon," J. Appl. Phys., 50, 874-880 (Feb. 1979).

65. L. Lundkvist, L. Lundstrom and C. Svensson, "Discharge of MNOS Structures," Solid-State Electron., 16, 811-823 (1973).

66. T. Katsube, Y. Adachi and T. Ikoma, "Memory Traps in MNOS Diodes Measured by Thermally Stimulated Currents," Solid-State Electron., **19**, 11-16 (1976).

67. L. Lundkvist, C. Svensson and B. Hansson, "Discharge of MNOS Structures at Elevated Temperatures," Solid-State Electron., **19**, 221-227 (1976).

68. K. Lehovec and A. Fedotowsky, "Charge Retention of MNOS Devices Limited by Frenkel-Poole Detrapping," Appl. Phys. Lett., **32**, 335-338 (1 Mar. 1978).

69. K. Lehovec and A. Fedotowsky, "Charge Centroid in MNOS Devices," J. Appl. Phys., 48, 2955-2960 (July 1977).

70. J.A. Topich and E.T. Yon, "The Effects of High Temperature Annealing on MNOS Devices," J. Electrochem. Soc., 123, 535-539 (Apr. 1976).

71. V.J. Kapoor and S.B. Bibyk, "Energy Distribution of Electron-Trapping Centers in Low Pressure Chemically Vapor-Deposited Si_3N_4 Films," Thin Solid Films, 78, 193-201 (1981).

72. M.H. White and J.R. Cricchi, "Characterization of Thin-Oxide MNOS Memory Transistors," IEEE Trans. Electron. Devices, **ED-19**, 1280-1288 (Dec. 1972).

73. L.I. Popova, P.K. Vitanov and B.Z. Antov, "Interface States in MNOS Systems," Thin Solid Films, 51, 305-309 (1978).

74. Y.C. Sun and C.T. Sah, "Interface Edge Effect and Its Contribution to the Frequency Dispersion of Metal-Oxide-Semiconductor Admittance," Solid-State Electron., **24**, 569-576 (1981).

75. N.M. Johnson, D.K. Biegelsen, M.D. Moyer, S.T. Chang, E.H. Poindexter and P.J. Caplan, "Characteristic Electronic Defects at the Si-SiO_2 Interface," Appl. Phys. Lett., **43**, 563-565 (15 Sep. 1983).

76. K.O. Jeppson and C.M. Svensson, "Negative Bias Stress of MOS Devices at High Electric Fields and Degradation of MNOS Devices," J. Appl. Phys., **48**, 2004-2014 (May 1977).

77. T.J. Tredwell and C.R. Viswanathan, "Determination of Interface-State Parameters in a MOS Capacitor by DLTS," Solid-State Electron., **23**, 1171-1178 (1980).

78. H. Flietner and Ngo Duong Sinh, "Character of Si-SiO_2 Interface States from Analysis of CV Term Spectra," Phys. Status Solidi (a), **37**, 533-539 (1976).

79. N.M. Johnson, D.J. Bartelink and M. Schulz, "Transient Capacitance Measurements of Electronic States at the SiO_2-Si Interface," Proc. Intern. Topical Conf. on the Physics of SiO_2 and Its Interfaces, 421-427, Pergamon, New York (1978).

80. E. Arnold and H. Schauer, "Measurements of Interface State Density in MNOS Structures," Appl. Phys. Lett., **32**, 333-335 (1 Mar. 1978).

81. H-C.W. Huang, "Generation of Interface States Due to Emission of Leakage Electrons from Silicon Substrate into Silicon Dioxide," Appl. Phys. Lett., 30, 533-535 (15 May 1977).

82. H. Schauer, E. Arnold and P.C. Murau, "Interface States and Memory Decay in MNOS Capacitors," IEEE Trans. Electron. Devices, **ED-25**, 1037-1041 (Aug. 1978).

83. A. Sher, H.J. Hoffman, P. Su and Y.H. Tsuo, "Techniques for Improving the Si-SiO_2 Interface Characterization," J. Appl. Phys., **54**, 5183-5198 (Sep. 1983).

84. M. Jourdain, G. Salace, C. Petit, M. Favre, J. Despujols and V. Le Goascoz, "Effect of High Field Stresses on Interface States of n-MOS Capacitors," Solid-State Electron., **26**, 251-257 (1983).

85. L.S. Wei and J.G. Simmons, "Trapping, Emission and Generation in MNOS Memory Devices," Solid-State Electron., **17**, 591-598 (1974).

86. H.A. Mar and J.G. Simmons, "A Review of the Techniques Used to Determine Trap Parameters in the MNOS Structure," IEEE Trans. Electron. Devices, **ED-24**, 540-546 (May 1977).

87. P. Vitanov and I. Eisele, "A New Transient Capacitance Technique for the Determination of Interface Traps in Metal-Oxide-Semiconductor Transistors," J. Appl. Phys., **54**, 5227-5230 (Sep. 1983).

88. H. Lu, "Investigation of Interface States and Failure Mechanisms in MNOS Devices," Ph.D. Dissertation, University of Florida (1984).

89. J.S. Uranwala, J.G. Simmons and H.A. Mar, "The Determination of the Energy Distribution of Interface Traps in Metal-Nitride-Oxide-Silicon (Memory) Devices Using Non-Steady-State Techniques," Solid-State Electron., **19**, 375-380 (1976).

90. P.V. Gray and D.M. Brown, "Density of SiO_2-Si Interface States," Appl. Phys. Lett., **8**, 31-33 (15 Jan. 1966).

91. K.L. Wang, "MOS Interface-State Density Measurements Using Transient Capacitance Spectroscopy," IEEE Trans. Electron. Devices, ED-**27**, 2231-2239 (Dec. 1980).

92. R.F. Pierret, "Photo-Thermal Probing of Si-SiO_2 Surface Centers - I, Theory," Solid-State Electron., **19**, 577-591 (1976).

93. M. Kamada and A. Yagi, "Effects of Silicon-Nitride Films on the Electronic Properties of Si-SiO_2 Interfaces," Japan. J. Appl. Phys., **21**, 1684-1688 (Dec. 1982).

94. T. Katsube, K. Kakimoto and T. Ikoma, "Temperature and Energy Dependences of Capture Cross Sections at Surface States in Si Metal-Oxide-Semiconductor Diodes Measured by Deep Level Transient Spectroscopy," J. Appl. Phys., **52**, 3504-3508 (May 1981).

95. J.D. Joannopoulos, "Theory of Fluctuations and Localized States in Amorphous Tetrahedrally Bonded Solids," Phys. Rev. B, 16, 2764-2780 (15 Sep. 1977).

96. R.B. Laughlin, J.D. Joannopoulos and D.J. Chadi, "Electronic States of Si-SiO_2 Interfaces," Proc. Intern. Topical Conf. on the Physics of SiO_2 and Its Interfaces, 321-327, Pergamon, New York (1978).

97. H.J. Stein, P.S. Peercy and R.J. Sokel, "Post-Deposition High Temperature Processing of Silicon Nitride," Thin Solid Films, **101**, 291-298 (1983).

98. H.E. Maes and G.L. Heyns, "Influence of a High-Temperature Hydrogen Anneal on the Memory Characteristics of p-channel MNOS Transistors," J. Appl. Phys., **51**, 2706-2713 (May 1980).

99. H.E. Maes and R.J. Van Overstraeten, "Memory Loss in MNOS Capacitors," J. Appl. Phys., **47**, 667-671 (Feb. 1976).

100. R.A. Williams and M.M.E. Beguwala, "The Effect of Electrical Conduction of Si_3N_4 on the Discharge of MNOS Memory Transistors," IEEE Trans. Electron. Devices, **ED-25**, 1019-1023 (Aug. 1978).

101. N. Endo, "Charge Distributions in Silicon Nitride of MNOS Devices," Solid-State Electron., **21**, 1153-1156 (1978).

102. C.S. Dobbs, W.D. Brown and J.R. Yeargan, "Charge Loss in Metal-Nitride-Oxide-Semiconductor (MNOS) Devices at High Temperatures," Solid-State Electron., **26**, 427-432 (1983).

103. R.V. Jones and W.D. Brown, "Endurance and Retention of MNOS Devices over the Temperature Range from -50°C to +125°C," J. Electron. Mat., **10,#6**, 959-972 (1981).

104. W.D. Brown, "Retention and Endurance Characteristics of HCl-Annealed and Unannealed MNOS Capacitors," Solid-State Electron., **22**, 373-378 (1979).

105. C.A. Neugebauer and J.F. Burgess, "Endurance and Memory Decay of MNOS Devices," J. Appl. Phys., **47**, 3182-3191 (July 1976).

106. F.L. Schuermeyer and C.R. Young, "Endurance Studies on MNOS Devices," J. Appl. Phys., 49, 4556-4559 (Aug. 1978).

107. H. Bernt and J.W. Scholtens, "Interpretation of Flatband Voltage Shifts in Terms of Charge Distributions in MNOS Structures," Solid-State Electron., 25, 843-850 (1982).

108. M.H. White, J.W. Dzimianski and M.C. Peckerar, "Endurance of Thin-Oxide Nonvolatile MNOS Memory Transistors," IEEE Trans. Electron. Devices, **ED-24**, 577-580 (May 1977).

109. V.I. Belyi, F.A. Kuznetsov, T.P. Smirnova, L.V. Chramova and L.Kh. Kravchenko, "Chemical Non-uniformity of Thin Dielectric Films Produced by Ammonolysis of Monosilane," Thin Solid Films, **37**, L39-L42 (1976).

110. R. Gale, F.J. Feigl, C.W. Magee and D.R. Young, "Hydrogen Migration under Avalanche Injection of Electrons in Si Metal-Oxide-Semiconductor Capacitors," J. Appl. Phys., **54**, 6938-6942 (Dec. 1983).

111. J.K. Wu, S.A Lyon and W.C. Johnson, "Temperature and Field Dependence of the Generation of Interface States in the Si-SiO_2 System after High-Field Stress," Appl. Phys. Lett., **42**, 585-587 (1 Apr. 1983).

112. R. W. Pryor, "A Mechanism for Endurance Failure in Metal-Nitride-Oxide-Semiconductor Device Structures," J. Appl. Phys., 52, 3702-3704 (May 1981).

APPENDIX A
PHYSICAL CONSTANTS

Electronic charge, $q = 1.602\text{x}10^{-19}$ C

Permittivity of free space, $\varepsilon_0 = 8.854\text{x}10^{-14}$ F/cm

Boltzmann constant, $k = 1.38\text{x}10^{-23}$ J/^{o}K

The following values are at 300^{o}K:

Thermal voltage, $1/\beta = kT/q = 0.02584$ eV

Relative permittivity, ε_r - Si: 11.9, SiO_2: 3.9, Si_3N_4: 7.5

Bandgap energy of silicon, $E_g = 1.107$ eV

Intrinsic carrier density of silicon, $n_i = 1.45\text{x}10^{10}$ cm^{-3}

Thermal velocity of electrons, $v_{th} = 10^7$ cm/s

Breakdown Field, F_{bd} (V/cm) - Si: $3\text{x}10^5$, SiO_2: $6\text{x}10^6$, Si_3N_4: 10^7

Density (g/cc) - Si: 2.33, SiO_2: 2.2, Si_3N_4: 3.44

APPENDIX B
MNOS QUASI-STATIC ANALYSIS PROGRAM

The MNOS quasi-static analysis program is listed below. The input parameters A1, CIN, CIP, CMN, ND, N, XNC, XNA and XND denote respectively the area of the n-region (cm^2), the insulator capacitance above the n-region (pF), the insulator capacitance above the p^+-grid (pF), the minimum measured capacitance (pF), the device number, the number of data points, the number of cycles, the acceptor density of the p^+-grid (cm^{-3}) and the donor density of the n-region (cm^{-3}). The input variables V, CMO and CM are respectively the bias voltage (V), the quasi-static measured capacitance (pF) and the high frequency measured capacitance (pF). The output variables V, PS, CSC, CPN, CPO, E and NSS are respectively the bias voltage (V), the surface potential (V), the space charge capacitance (pF), the p^+-n junction capacitance (pF), the quasi-static interface-state capacitance (pF), the energy (eV) and the interface-state density ($eV^{-1}cm^{-2}$).

```
C  MNOS QUASI-STATIC ANALYSIS PROGRAM.
C
      DIMENSION   CM(50),CMO(50),CS(50),E(50),PS(50),V(50),XNS(50)
C  INPUTS.
      READ 100, A1,CIN,CIP,CMN,ND,N,XNC,XNA,XND
      READ 110, (V(I), I=1,N)
      READ 120, (CMO(I), I=1,N)
      READ 120, (CM(I), I=1,N)
```

```
C  SCALING OF CAPACITANCES.
      DO 5 I=1,N
         CM(I)=1.E-12*CM(I)
    5    CMO(I)=1.E-12*CMO(I)
      CIN=1.E-12*CIN
      CIP=1.E-12*CIP
      CMN=1.E-12*CMN
      CMN=1./(1./(CMN-CIP)-1./CIN)
C  INITIALIZATION OF FLAGS.
      K=0
      L=0
      M=0
C  RECURSIVE DETERMINATION OF DONOR DENSITY AND BUILT-IN POTENTIAL.
      A=1.808E-16*A1*SQRT(XND)
      AA=A
      AD=A
   10 CA=1.
      PSD=.1
      XND=(A/A1)**2/3.267E-30
      D=2.103E20/XND**2
      VB1=.02584*ALOG(XNA*XND/2.103E20)
      VB=VB1/.02584-2.
      PS1=-VB*.02584-.9*PSD
   15 PS1=PS1+PSD
      H=PS1/.02584
      C=ABS(A*(D*(1.-EXP(-H))+EXP(H)-1.)/SQRT(D*(EXP(-H)+H-1.)
     *      +EXP(H)-H-1.))+A/SQRT(VB+H)
      IF  (C.GE.CA) GO TO 20
      CA=C
      GO TO 15
   20 PS1=PS1-2.*PSD
      PSD=.1*PSD
      IF (PSD.LT..0001) GO TO 25
      CA=1.
      GO TO 15
   25 IF (C.EQ.CMN) GO TO 40
      IF (C.GT.CMN) GO TO 30
      A=A+AD
      L=1
      GO TO 10
   30 IF (L.EQ.1) GO TO 35
      PRINT 270
      PRINT 200
      STOP
   35 A=A-AD
      AD=.1*AD
      IF (AD.GE..001*AA) GO TO 10
```

```
C  DETERMINATION OF SURFACE POTENTIAL, SPACE CHARGE CAPACITANCE,
C  P-N JUNCTION CAPACITANCE, INTERFACE-STATE CAPACITANCE, ENERGY
C  AND INTERFACE-STATE DENSITY.
   40 B=1.602E-19*A1
      L=0
      PB=.02584*ALOG(XND/1.45E10)
      B1=.5535+PB
      PS1=.2
   45 DO 60 I=1,N
         I1=I
         IF (K.EQ.1) I1=N-I+1
         IF (I1.EQ.L) GO TO 80
         CS1=1./(1./(CM(I1)-CIP)-1./CIN)
         PSD=.1
         IF (K.EQ.1) PSD=-.1
         IF (K.EQ.1.OR.I.EQ.1) GO TO 50
         IF (CS1.GE.CSP) GO TO 55
   50    CALL SPOT (A,CPN,CS1,D,M,PSD,PS1,VB)
         CSP=CS1
         PS(I1)=PS1
         E(I1)=PS1+B1
         CMO(I1)=1./(1./(CMO(I1)-CIP)-1./CIN)-CS1
         XNS(I1)=CMO(I1)/B
         CM(I1)=CPN
         CS(I1)=CS1-CPN
         GO TO 60
   55    K=1
         L=I-1
         M=0
         PS1=-VB*.02584-.09999
         GO TO 70
   60 CONTINUE
      GO TO 80
   70 GO TO 45
C  OUTPUTS.
   80 PRINT 200
      PRINT 210
      PRINT 220, ND
      PRINT 230, XNC
      PRINT 240, PB
      PRINT 250, VB1
      PRINT 260, XND
      PRINT 300
      DO 90 I=1,N
   90    PRINT 310, V(I), PS(I), CS(I)*1.E12, CM(I)*1.E12,
     *               CMO(I)*1.E12, E(I), XNS(I)
      PRINT 200
      STOP
```

```
  100 FORMAT (E6.3,3F3.1,I3,I2,3E5.1)
  110 FORMAT (20F4.2)
  120 FORMAT (20F3.1)
  200 FORMAT ('1')
  210 FORMAT (///T10,'MNOS QUASI-STATIC ANALYSIS.     (AAMD)')
  220 FORMAT (///T10,'DEVICE # : ',I3)
  230 FORMAT (/T10,'# OF CYCLES = ',E8.2)
  240 FORMAT (/T10,'BULK POTENTIAL = ',F5.3,' V')
  250 FORMAT (/T10,'BUILT-IN POTENTIAL = ',F5.3,' V')
  260 FORMAT (/T10,'EFFECTIVE DONOR DENSITY = ',E10.4,' /C.CM')
  270 FORMAT (///T10,'MIN. CAP. EXCEEDED INITIALLY.')
  300 FORMAT (/////T5,'V(V)',T15,'PS(V)',T25,'CSC(PF)',T35,'CPN(PF)',
     *        T45,'CPO(PF)',T60,'E(EV)',T70,'NSS(/EV/SQ.CM)'/)
  310 FORMAT (T5,F5.2,T15,F6.3,T25,F6.1,T35,F6.1,T45,E9.3,T60,F5.3,
     *        T70,E9.3)
      END
C
C
C
      SUBROUTINE SPOT (A,CPN,CS1,D,M,PSD,PS1,VB)
C
C  DETERMINATION OF SURFACE POTENTIAL, SPACE CHARGE CAPACITANCE AND
C  P-N JUNCTION CAPACITANCE.
      IF ((M.EQ.1) GO TO 40
      C=1.
   10 PS1=PS1-PSD
   20 CA=C
   30 IF (ABS(PS1).LT..0005) PS1=.0004
      H=PS1/.02584
      CPN=A/SQRT(VB+H)
      C=ABS(A*(D*(1.-EXP(-H))+EXP(H)-1.)/SQRT(D*(EXP(-H)+H-1.)
     *   +EXP(H)-H-1.))+CPN
      IF (C.EQ.CS1) GO TO 70
      IF (C.GE.CA.AND.PSD.GT..002) GO TO 50
      IF (C.GE.CA) GO TO 60
      IF (C.GT.CS1) GO TO 10
      PS1=PS1+PSD
   40 PSD=.1*PSD
      IF (ABS(PSD).LT..0002) GO TO 70
      C=1.
      GO TO 20
   50 PS1=PS1+.9*PSD
      PSD=.1*PSD
      GO TO 30
   60 PS1=PS1+PSD
   70 M=1
      RETURN
      END
```

APPENDIX C
MNOS C-V CURVE SIMULATION PROGRAM

The MNOS C-V curve simulation program for the model in Figure 4.7 is listed below. The input parameters A1, CIN, CIP, N, XNA and XND have the same notations as in Appendix B, while RS and NF represent the series resistance (Ohms) and the number of frequencies (including the quasi-static case) respectively. The input variables E, F, R, SN and XNS are the interface-state energy (eV), the frequency (Hz), the reciprocal of the space charge loss conductance (Ohms), the majority carrier capture cross-section of a continuum of interface states (cm^2) and the interface-state density ($eV^{-1}cm^{-2}$) respectively. The output variables PS and CM have the same notations as in Appendix B.

```
C  MNOS C-V CURVE SIMULATION PROGRAM
C
      COMPLEX  Y
      DIMENSION CPN(50),CSO(50),E(50),F(30),R(30),SN(50),XNS(50)
C  INPUTS.
      READ 100, A1,CIN,CIP,RS,N,NF,XNA,XND
      READ 110, (E(I), I=1,N)
      READ 120, (F(I), I=1,NF)
      READ 130, (R(I), I=1,NF)
      READ 140, (SN(I), I=1,N)
      READ 150, (XNS(I), I=1,N)
C  COMPUTATION OF CONSTANTS.
      CIN=1.E-12*CIN
      CIP=1.E-12*CIP
      A=1.808E-15*A1*SQRT(XND)
      A1=1.602E-19*A1
      D=2.103E20/XND**2
      PB=.02584*ALOG(XND/1.45E10)+.5535
      VB=ALOG(XNA*XND/2.103E20)
```

```
C  COMPUTATION OF BIAS DEPENDENT VARIABLES.
      DO 10 I=1,N
         E(I)=E(I)-PB
         H=E(I)/.02584
         CPN(I)=A/SQRT(VB+H-2.)
         CSO(I)=A*ABS((D*(1.-EXP(-H))+EXP(H)-1.)/SQRT(D*(EXP(-H)
     *            +H-1.)+EXP(H)-H-1.))
         SN(I)=1./(1.E7*SN(I)*XND*EXP(H))
   10    XNS(I)=A1*XNS(I)
C  COMPUTATION OF FREQUENCY AND BIAS DEPENDENT VARIABLES; OUTPUTS.
      PRINT 200
      PRINT 210
      DO 50 I=1,NF
         PRINT 220, F(I)
         F(I)=6.283*F(I)
         PRINT 230
         DO 40 J=1,N
            T=F(I)*SN(J)
            IF (F(I).LT.1.) GO TO 20
            CP=XNS(J)*ATAN(T)/T
            GP=XNS(J)*ALOG(1.+T**2)*F(I)/(2.*T)
            GO TO 30
   20       CM=(CIP+1./(1./(CSO(J)+CPN(J)+XNS(J))+1./CIN))*1.E12
            GO TO 40
   30       Y=1./(1./(1./(R(I)+1./CMPLX(0.,F(I)*CSO(J)))+CMPLX(GP,
     *        F(I)*(CPN(J)+CP)))+1./CMPLX(0.,F(I)*CIN))
            Y=1./(1./(Y+CMPLX(0.,F(I)*CIP))+RS)
            CM=AIMAG(Y)*1.E12/F(I)
   40       PRINT 240, E(J), CM
   50 CONTINUE
      PRINT 200
      STOP
  100 FORMAT (E6.3,2F3.1,F3.0,2I2,E5.1,E7.3)
  110 FORMAT (40F2.2)
  120 FORMAT (20E4.1)
  130 FORMAT (15E5.2)
  140 FORMAT (15E5.1)
  150 FORMAT (12E6.2)
  200 FORMAT ('1')
  210 FORMAT (///T10,'MNOS MODEL VERIFICATION.     (AAMD)')
  220 FORMAT (///T10,'FREQUENCY = ',E8.2//)
  230 FORMAT (T15,'PS (V)',T25,'CM (PF)'/)
  240 FORMAT (T15,F6.3,T25,F5.1)
      END
```

BIOGRAPHICAL SKETCH

Marino de Almeida was born on July 26, 1947, in Colombo, Sri Lanka. He attended St. Joseph's College, Colombo, and Aquinas College before entering the University of Sri Lanka, Katubedde, for his undergraduate studies. He obtained the B.Sc.Eng. degree in electronic engineering in November, 1972, and served as an instructor for ten months at his undergraduate alma mater. He joined the Posts & Telecommunications Department of Sri Lanka as an engineer in October, 1973. He came to the U.S. in June, 1979, to attend the University of Florida, where he obtained the M.E. degree in June, 1981.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sheng S. Li, Chairman
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Dorothea Burk

Dorothea Burk
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gijs Bosman
Assistant Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul H Holloway
Paul Holloway
Professor of Materials Science and Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arun Kumar Varma
Arun Varma
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1986

Shafer
Dean, College of Engineering

Dean, Graduate School